AN

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 June 2003 (26.06.2003)

PCT

(10) International Publication Number
WO 03/052049 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US02/21767

(22) International Filing Date: 9 July 2002 (09.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/304,298	9 July 2001 (09.07.2001)	US
60/305,324	13 July 2001 (13.07.2001)	US
60/307,003	19 July 2001 (19.07.2001)	US
60/308,185	27 July 2001 (27.07.2001)	US
60/310,096	3 August 2001 (03.08.2001)	US
60/311,551	10 August 2001 (10.08.2001)	US
60/363,649	8 March 2002 (08.03.2002)	US

(71) Applicant *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **THORNTON, Michael** [US/US]; 9 Medway Road, Woodside, CA 94062 (US). **AU-YOUNG, Janice, K.** [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). **AZIMZAI, Yalda** [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US). **BANDMAN, Olga** [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). **BARROSO, Ines** [PT/GB]; 38 Eden Street, Cambridge CB1 1EL (GB). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **BECHA, Shanya, D.** [US/US]; 21062 Gary Drive # 117, Castro Valley, CA 94546 (US). **BOROWSKY, Mark, L.** [US/US]; 122 Orchard Avenue, Redwood City, CA 94061 (US). **DING, Li** [CN/US]; 11036 Ridgecorde Lane, Creve Coeur, MO 63141 (US). **DUGGAN, Brendan, M.** [AU/US]; 243 Buena Vista Avenue # 306, Sunnyvale, CA 94086 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **EMERLING, Brooke, M.** [US/US]; 1735 Woodland Avenue # 71, Palo Alto, CA 94303 (US). **FORSYTHE, Ian, J.** [US/US]; 308 Roble Avenue, Redwood City, CA 94061 (US). **GANDHI, Ameena, R.** [US/US]; 705 5th Avenue, San Francisco, CA 94118 (US). **GIETZEN, Kimberly, J.** [US/US]; 691 Los Huecos Drive, San Jose, CA 95123 (US). **GORVAD, Ann, E.** [US/US]; 369 Marie Common, Livermore, CA 94550 (US). **GRIFFIN, Jennifer, A.** [US/US]; 33691 Mello Way, Fremont, CA 94555 (US). **GURURAJAN, Rajagopal** [IN/US]; 5591 Dent Avenue, San Jose, CA 95118 (US). **HAFALIA, April, J.A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **RING, Huijun, Z.** [US/US]; 625 Orange Avenue, Los Altos, CA 94022 (US). **ISON, Craig, H.** [US/US]; 1242 Weathersfield Way, San Jose, CA 95118 (US). **JONES, Karen, Anne** [GB/GB]; 2 Castle Court, Saffron Walden, Essex, Greater London CB10 1BH (GB). **LAL, Preeti, G.** [US/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **LEE, Ernestine, A.** [US/US]; 20523 Crow Creek Road, Castro Valley, CA 94552 (US). **LEE, Sally** [US/US]; 3643 26th Street, San Francisco, CA 94110 (US). **LI, Joana, X.** [US/US]; 1264 Geneva Avenue, San Francisco, CA 94112 (US). **LU, Dyung Aina, M.** [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). **MARQUIS, Joseph, P.** [US/US]; 4428 Lazy Lane, San Jose, CA 95135 (US). **LEHR-MASON, Patricia, M.** [US/US]; 360 Clarke Lane, Morgan Hill, CA 95014 (US). **CHAWLA, Narinder, K.** [US/US]; 33 Union Square, #712, Union City, CA 94587 (US). **ARVIZU, Chandra, S.** [US/US]; 1706 Morocco Drive, San Jose, CA 95125 (US). **SANJANWALA, Bharati** [IN/US]; 210 Silvia Court, Los Altos, CA 94024 (US). **SORNASSE, Thierry** [BE/US]; 1525 Fordham Way, Mountain View, CA 94040 (US). **SWARNAKAR, Anita** [CA/US]; 8 Locksley Avenue #5D, San Francisco, CA 94122 (US). **TANG, Y. Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **THANGAVELU, Kavitha** [IN/US]; 1950 Montecito Avenue #23, Mountain View, CA 94043 (US). **TRAN, Bao** [US/US]; 750 Salberg Avenue, Santa Clara, CA 95051 (US). **TRAN, Uyen, K.** [US/US]; 2638 Mabury Square, San Jose, CA 95133 (US). **WARREN, Bridget, A.** [US/US]; 2250 Homestead Court # 2, Los Altos, CA 94024 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). **YAO, Monique, G.** [US/US]; 1189 Woodgate Drive, Carmel, IN 46033 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **YUE, Huibin** [US/US]; 1170 South Stelling Road, Cupertino, CA 95014 (US). **ZEBARJADIAN, Yeganeh** [IR/US]; 830 Junipero Serra Boulevard, San Francisco, CA 94127 (US). **CHANG, Hsin-Ru** [US/US]; 326 Treasure Island Drive, Belmont, CA 94002 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

[Continued on next page]

(54) Title: MOLECULES FOR DISEASE DETECTION AND TREATMENT

(57) Abstract: Various embodiments of the invention provide human molecules for disease detection and treatment (MDDT) and polynucleotides which identify and encode MDDT. Embodiments of the invention also provide expression vectors, host cells, antibodies, agonists, and antagonists. Other embodiments provide methods for diagnosing, treating, or preventing disorders associated with aberrant expression of MDDT.

WO 03/052049 A2

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

MOLECULES FOR DISEASE DETECTION AND TREATMENT

TECHNICAL FIELD

The invention relates to novel nucleic acids, molecules for disease detection and treatment encoded by these nucleic acids, and to the use of these nucleic acids and proteins in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders, diseases treated with steroids and disorders caused by the metabolic response to treatment with steroids. The invention also relates to the assessment of the effects of exogenous compounds on the expression of nucleic acids and molecules for disease detection and treatment.

BACKGROUND OF THE INVENTION

It is estimated that only 2% of mammalian DNA encodes proteins, and only a small fraction of the genes that encode proteins are actually expressed in a particular cell at any time. The various types of cells in a multicellular organism differ dramatically both in structure and function, and the identity of a particular cell is conferred by its unique pattern of gene expression. In addition, different cell types express overlapping but distinctive sets of genes throughout development. Cell growth and proliferation, cell differentiation, the immune response, apoptosis, and other processes that contribute to organismal development and survival are governed by regulation of gene expression. An example of a mammalian apoptosis-associated protein is Diablo, which can bind to apoptosis inhibition proteins and antagonize their antiapoptotic effect, a function analogous to that of the proapoptotic Drosophila molecules, Grim, Reaper, and HID (Ekert, P.G. et al. (2001) J. Cell Biol. 152:483-90). Appropriate gene regulation also ensures that cells function efficiently by expressing only those genes whose functions are required at a given time. Factors that influence gene expression include extracellular signals that mediate cell-cell communication and coordinate the activities of different cell types. Gene expression is regulated at the level of DNA and RNA transcription, and at the level of mRNA translation.

Aberrant expression or mutations in genes and their products may cause, or increase susceptibility to, a variety of human diseases such as cancer and other cell proliferative disorders. The identification of these genes and their products is the basis of an ever-expanding effort to find markers for early detection of diseases and targets for their prevention and treatment. For example, cancer represents a type of cell proliferative disorder that affects nearly every tissue in the body. The development of cancer, or oncogenesis, is often correlated with the conversion of a normal gene into a cancer-causing gene, or oncogene, through abnormal expression or mutation. Oncoproteins, the products of oncogenes, include a variety of molecules that influence cell proliferation, such as growth factors, growth factor receptors, intracellular signal transducers, nuclear transcription factors, and

cell-cycle control proteins. In contrast, tumor-suppressor genes are involved in inhibiting cell proliferation. Mutations which reduce or abrogate the function of tumor-suppressor genes result in aberrant cell proliferation and cancer. Thus a wide variety of genes and their products have been found that are associated with cell proliferative disorders such as cancer, but many more may exist that are yet to be discovered.

DNA-based arrays can provide an efficient, high-throughput method to examine gene expression and genetic variability. For example, SNPs, or single nucleotide polymorphisms, are the most common type of human genetic variation. DNA-based arrays can dramatically accelerate the discovery of SNPs in hundreds and even thousands of genes. Likewise, such arrays can be used for SNP genotyping in which DNA samples from individuals or populations are assayed for the presence of selected SNPs. These approaches will ultimately lead to the systematic identification of all genetic variations in the human genome and the correlation of certain genetic variations with disease susceptibility, responsiveness to drug treatments, and other medically relevant information. (See, for example, Wang, D.G. et al. (1998) Science 280:1077-1082.)

DNA-based arrays can also provide a simple way to explore the expression of a single polymorphic gene. When the expression of a single gene is explored, DNA-based arrays are employed to detect the expression of specific gene variants. For example, a p53 tumor suppressor gene array is used to determine whether individuals are carrying mutations that predispose them to cancer. A cytochrome p450 gene array is useful to determine whether individuals have one of a number of specific mutations that could result in increased drug metabolism, drug resistance or drug toxicity.

DNA-based array technology is especially important for the rapid analysis of global gene expression patterns. There is a growing awareness that gene expression is affected in a global fashion. In some cases the interactions may be expected, such as when the genes are part of the same signaling pathway. In other cases, such as when the genes participate in separate signaling pathways, the interactions may be totally unexpected. Therefore, DNA-based arrays can be used to investigate how genetic predisposition, disease, or therapeutic treatment affects the expression of a large number of genes. In this case, it is useful to develop a profile, or transcript image, of all the genes that are expressed and the levels at which they are expressed in that particular tissue. A profile generated from an individual or population affected with a certain disease or undergoing a particular therapy may be compared with a profile generated from a control individual or population. Such analysis does not require knowledge of gene function, as the expression profiles can be subjected to mathematical analyses which simply treat each gene as a marker. Furthermore, gene expression profiles may help dissect biological pathways by identifying all the genes expressed, for example, at a certain developmental stage, in a particular tissue, or in response to disease or treatment. (See, for

example, Lander, E.S. et al. (1996) Science 274:536-539.)

Certain genes are known to be associated with diseases because of their chromosomal location, such as the genes in the myotonic dystrophy (DM) regions of mouse and human. The mutation underlying DM has been localized to a gene encoding the DM-kinase protein, but another active gene, DMR-N9, is in close proximity to the DM-kinase gene (Jansen, G. et al. (1992) Nat. Genet. 1:261-266). DMR-N9 encodes a 650 amino acid protein that contains WD repeats, motifs found in cell signaling proteins. DMR-N9 is expressed in all neural tissues and in the testis, suggesting a role for DMR-N9 in the manifestation of mental and testicular symptoms in severe cases of DM (Jansen, G. et al. (1995) Hum. Mol. Genet. 4:843-852).

Other genes are identified based upon their expression patterns or association with disease syndromes. For example, autoantibodies to subcellular organelles are found in patients with systemic rheumatic diseases. A recently identified protein, golgin-67, belongs to a family of Golgi autoantigens having alpha-helical coiled-coil domains (Eystathioy, T. et al. (2000) J. Autoimmun. 14:179-187). The Stac gene was identified as a brain specific, developmentally regulated gene. The Stac protein contains an SH3 domain, and is thought to be involved in neuron-specific signal transduction (Suzuki, H. et al. (1996) Biochem. Biophys. Res. Commun. 229:902-909).

Evi-5 is a site of retroviral integration in AKXD T-cell lymphoma cells. Tumors with Evi-5 integrations have also been shown to possess other integration sites associated with T-cell disease. Retroviral disease induction occurs as a result of insertional mutagenesis of cellular proto-oncogenes or tumor suppressor genes. The AKXD recombinant inbred murine model is useful in the study of retrovirally-induced myeloid tumors, as well as T- and B-cell leukemias (Liao, X. et al. (1997) Oncogene 14:1023-1029). Lymphomas with integrations in Evi-5 may also possess integrations in *Myc*, in sites located near and activating *Myc*, or that synergize with *Myc*. This suggests a possible cooperation between Evi-5 with *Myc* in tumor induction, consistent with other observations showing that *Myc* is a frequent target of retroviral integration in mouse and rat T-cell lymphomas.

The contiguous gene deletion syndrome AMME is characterized by Alport syndrome, midface hypoplasia, mental retardation and elliptocytosis and is caused by a deletion in Xq22.3, comprising several genes including COL4A5, FACL4 and AMMECR1. AMMECR1, found in eukaryotic and prokaryotic cells, contains six exons and codes for a protein with a molecular mass of 35.5 kDa. Evidence suggests that this protein is a regulatory factor potentially involved in the development of AMME contiguous gene deletion syndrome. The mouse ortholog has 95.2% identity at the amino acid level with human AMMECR1 and maps to chromosome MmuXF1-F3 (Vitelli, F. et al. (1999) Genomics 55:335-340; Vitelli, F. et al. (2000) Cytogenet. Cell Genet. 88:259-263).

Sporulation-induced transcript 4 (SIT4) gene is a type 2A-related serine/threonine protein phosphatase which when overexpressed confers lithium tolerance in galactose medium to the budding

yeast, Saccharomyces cerevisiae. It is a regulator of the cell cycle and is involved in nitrogen sensing, normal g1 cyclin expression, and bud initiation (Masuda, C. A. et al. (2000) J. Biol. Chem. 275:30957-30961). The SIT4-associated proteins (SAPs), SAP155, SAP185, SAP190, and probably SAP4, associate with SIT4 in separate complexes. The SAPs are not functional in the absence of SIT4 and likewise, SIT4 is not functional in the absence of the SAPs. However, SAPs and SIT4 have distinct functions (Luke, M. M. et al. (1996) Mol. Cell. Biol.16:2744-2755). C11orf23 is a human ortholog of the yeast SAP family. C11orf23 has been mapped to the 400-kb region of the IDDM4 locus of chromosome 11q13, a region involved in type 1 diabetes (Twells, R. C. et al. (2001) Genomics72:231-242).

Dendritic cells are antigen-presenting cells that play a major role in the initial phases of the immune response. Dendritic cells located in peripheral tissues are generally immature and exhibit a strong capacity to capture surrounding antigens whereas they exhibit limited T cell activation capacity. Reciprocally, mature dendritic cells found in lymphoid organs exhibit a strong capacity to activate T cells but have lost most of their ability to pick up new antigens. Dendritic cells migrating out of transplanted organ play a major role in the induction of graft rejection. Therefore, genes that are modulated during the maturation of dendritic cells represent potential targets for drugs aimed at limiting the rejection of transplanted organs.

Rho-family GTPases are critical mediators of dendritic growth and remodeling. Three of these Rho GTPases, RhoA, Rac1 and Cdc42 (cell division cycle 42), regulate distinct aspects of dendritic development, such as dendrite initiation, dendrite growth, dendrite branching, and spine formation. In cortical neurons, Rho-family GTPases play a central role in determining the number of primary dendrites in both pyramidal and non-pyramidal neurons. Research suggests that Rac1 is an important effector of dendrite initiation and that a common effector of Rac1 and Cdc42 mediates dendrite initiation. Suggested effectors include the p21-activated kinase (PAK) family of serine threonine kinases and LIM-domain-containing protein kinase, which can modulate actin dynamics by phosphorylation of cofilin.

Rho-family GTPases also can influence large-scale dendritic remodeling. Many neurons in the cortex initially acquire a pyramidal morphology and undergo a developmentally-regulated transformation into non-pyramidal neurons. This transformation involves the withdrawal of the apical dendrite and the extension of new primary dendrites, and is inhibited by expression of dominant-negative Cdc42 and, to a lesser extent, by dominant-negative Rac1. This inhibition suggests that the acquisition of cell-type-specific dendritic morphologies is under the control of Rac1 and Cdc42 signaling Redmond, L. and Ghosh, A. (2001) Curr. Opin. Neurobiol. 11:111-117).

ADP-ribosylation factors (ARFs) are small guanine-nucleotide-binding proteins that regulate membrane traffic and organelle structure in eukaryotic cells. In general, the inactive GDP-bound

form of ARF is soluble, although it can associate weakly with membranes, whereas the active GTP-bound form binds tightly to the membrane. ARFs function on membrane surfaces where they encounter their effectors and regulators, the guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs). ARF effectors include lipid-modifying enzymes and cytosolic coat complexes (such as COPI) that are recruited onto membranes by ARF-GTP. Hence, ARF activation leads to changes in both the lipid and protein composition of the membrane on which it is localized; changes which in turn result in modulation of membrane structure and function.

ARF proteins are highly conserved and have been found in all eukaryotic organisms examined. Mammalian ARF proteins are divided into three classes: Class I (ARF1–ARF3), Class II (ARF4 and ARF5) and Class III (ARF6). Class I ARFs are involved in trafficking in the ER–Golgi and endosomal systems, and their functions have been extensively studied. ARF1 binding to endosomal membranes is regulated by endosomal pH, which explains the pH dependence of COPI binding to endosomes. The Class III ARF, ARF6, functions exclusively in the endosomal–plasma membrane system. ARF6 is involved in endosomal recycling to the plasma membrane (PM), in regulated secretion, and in coordinating actin cytoskeleton changes at the PM. ARF6 is present at the apical surface of Madin Darby Canine Kidney (MDCK) epithelial cells, where it plays a role in modulating clathrin endocytosis. ARF6 has also been implicated in Fc-mediated phagocytosis in macrophages and in insulin stimulation of adipsin secretion and Glut4 translocation. By contrast, virtually nothing is known about the functions of the class II ARFs (Donaldson, J. D. and Jackson, C. L. (2000) Curr. Opin. Cell Biol. 12:475-482).

Expression profiling

Microarrays are analytical tools used in bioanalysis. A microarray has a plurality of molecules spatially distributed over, and stably associated with, the surface of a solid support. Microarrays of polypeptides, polynucleotides, and/or antibodies have been developed and find use in a variety of applications, such as gene sequencing, monitoring gene expression, gene mapping, bacterial identification, drug discovery, and combinatorial chemistry.

One area in particular in which microarrays find use is in gene expression analysis. Array technology can provide a simple way to explore the expression of a single polymorphic gene or the expression profile of a large number of related or unrelated genes. When the expression of a single gene is examined, arrays are employed to detect the expression of a specific gene or its variants. When an expression profile is examined, arrays provide a platform for identifying genes that are tissue specific, are affected by a substance being tested in a toxicology assay, are part of a signaling cascade, carry out housekeeping functions, or are specifically related to a particular genetic predisposition, condition, disease, or disorder.

The potential application of gene expression profiling is particularly relevant to improving

diagnosis, prognosis, and treatment of disease. For example, both the levels and sequences expressed in tissues from subjects with a cardiovascular disorder may be compared with the levels and sequences expressed in normal tissue.

Atherosclerosis and the associated coronary artery disease and cerebral stroke represent the most common cause of death in industrialized nations. Although certain key risk factors have been identified, a full molecular characterization that elucidates the causes and provides care for this complex disease has not been achieved. Molecular characterization of growth and regression of atherosclerotic vascular lesions requires identification of the genes that contribute to features of the lesion including growth, stability, dissolution, rupture and, most lethally, induction of occlusive vessel thrombus. Vascular lesions principally involve the vascular endothelium and the surrounding smooth muscle tissue.

Development of atherosclerosis is understood to be induced by the presence of circulating lipoprotein. Lipoproteins, such as the cholesterol-rich low-density lipoprotein (LDL), accumulate in the extracellular space of the vascular intima, and undergo modification. Oxidation of LDL (Ox-LDL) occurs most avidly in the sub-endothelial space where circulating antioxidant defenses are less effective. Mononuclear phagocytes enter the intima, differentiate into macrophages, and ingest modified lipids including Ox-LDL. During Ox-LDL uptake, macrophages produce cytokines (e.g. tumor necrosis factor α (TNF-α) and interleukin-1 (IL-1)) and growth factors (e.g. M-CSF, VEGF, and PDGF-BB) that elicit further cellular events that modulate atherogenesis such as smooth muscle cell proliferation and production of extracellular matrix by vascular endothelium. Additionally, these macrophages may activate genes in endothelium and smooth muscle tissue involved in inflammation and tissue differentiation, including superoxide dismutatse (SOD), IL-8, and ICAM-1.

The vascular endothelium influences not only the three classically interacting components of hemostasis: the vessel, the blood platelets and the clotting and fibrinolytic systems of plasma, but also the natural sequelae: inflammation and tissue repair. Two principal modes of endothelial behavior may be differentiated, best defined as an anti- and a prothrombotic state. Under physiological conditions endothelium mediates vascular dilatation (formation of nitric oxide (NO), PGI_2, adenosine, hyperpolarising factor), prevents platelet adhesion and activation (production of adenosine, NO and PGI_2, removal of ADP), blocks thrombin formation (tissue factor pathway inhibitor, activation of protein C via thrombomodulin, activation of antithrombin III) and mitigates fibrin deposition (t- and scuplasminogen activator production). Adhesion and transmigration of inflammatory leukocytes are attenuated, e.g. by NO and IL-10, and oxygen radicals are efficiently scavenged (urate, NO, glutathione, SOD).

When the endothelium is physically disrupted or functionally perturbed by postischemic reperfusion, acute and chronic inflammation, atherosclerosis, diabetes and chronic arterial

hypertension, then completely opposing actions pertain. This prothrombotic, proinflammatory state is characterised by vaso-constriction, platelet and leukocyte activation and adhesion (externalization, expression and upregulation of, for example, von Willebrand factor, platelet activating factor, P-selectin, ICAM-1, IL-8, MCP-1, and TNF-α), promotion of thrombin formation, coagulation and fibrin deposition at the vascular wall (expression of tissue factor, PAI-1, and phosphatidyl serine) and, in platelet-leukocyte coaggregates, additional inflammatory interactions via attachment of platelet CD40-ligand to endothelial, monocyte and B-cell CD40. Since thrombin formation and inflammatory stimulation set the stage for later tissue repair, complete abolition of such endothelial responses cannot be the goal of clinical interventions aimed at limiting procoagulatory, prothrombotic actions of a dysfunctional vascular endothelium. (See, e.g., Becker et al. (2000) Z Kardiol 89:160-167.)

Tumor necrosis factor α is a pleiotropic cytokine that a mediates immune regulation and inflammatory responses. TNF-α-related cytokines generate partially overlapping cellular responses, including differentiation, proliferation, nuclear factor-κB (NF-κB) activation, and cell death, by triggering the aggregation of receptor monomers (Smith, C.A. et al. (1994) Cell 76:959-962). The cellular responses triggered by TNF-α are initiated through its interaction with distinct cell surface receptors (TNFRs). NF-κB is a transcription factor with a pivotal role in inducing genes involved in physiological processes as well as in the response to injury and infection. Activation of NF-κB involves the phosphorylation and subsequent degradation of an inhibitory protein, IKB,and many of the proximal kinases and adaptor molecules involved in this process have been elucidated. Additionally, the NF-κB activation pathway from cell membrane to nucleus for IL-1 and TNF-α is now understood (Bowie and O'Neill (2000) Biochem Pharmacol 59:13-23).

Monocyte chemoattractant protein-1 (MCP-1) is known to play an important role in the pathogenesis of atherosclerosis by inducing monocyte migration. TNF-α treatment of human umbilical vein endothelial cells (HUVECs) increased the cellular secretions of MCP-1 119-fold compared with untreated cells. Troglitazone, an insulin-sensitizing drug, significantly inhibited this TNF-α-induced increase in MCP-1 secretions and decreased mRNA levels (Ohta et al. (2000) Diabetes Res Clin Pract 48:171-176).

Treatment of confluent cultures of vascular smooth muscle cells (SMCs) with TNF-α suppresses the incorporation of [^{3}H]proline into both collagenase-digestible proteins (CDP) and noncollagenous proteins (NCP). Such suppression by TNF-α is not observed in confluent bovine aortic endothelial cells and human fibroblastic IMR-90 cells. TNF-α decreases the relative proportion of collagen types IV and V suggesting that TNF-α modulates collagen synthesis by SMCs depending on their cell density and therefore may modify formation of atherosclerotic lesions (Hiraga et al. (2000) Life Sci 66:235-244).

Human coronary artery smooth muscle cells (CASMC) are primary cells isolated from the tunica media (an intermediate muscular layer) of a human coronary artery. Vascular smooth muscle cells are a model of increasing significance in vascular biology. It is now well known that besides their obvious role in the regulation of vascular tone and, consequently, oxygen supply to various tissues, their behavior under inflammatory conditions is an important factor in the development of atherosclerosis and restenosis.

Human aortic endothelial cells (HAECs) are primary cells derived from the endothelium of a human aorta. HAECs have been used as an experimental model for investigating *in vitro* the role of the endothelium in human vascular biology. Activation of the vascular endothelium is considered to be a central event in a wide range of both physiological and pathophysiological processes, such as vascular tone regulation, coagulation and thrombosis, atherosclerosis, and inflammation.

Thus, vascular tissue genes differentially expressed during treatment of CASMC and HAEC cell cultures with TNFα may reasonably be expected to be markers of the atherosclerotic process.

The potential application of gene expression profiling is particularly relevant to improving diagnosis, prognosis, and treatment of disease. For example, both the levels and sequences expressed in tissues from subjects with ovarian cancer may be compared with the levels and sequences expressed in normal tissue. Ovarian cancer is the leading cause of death from a gynecologic cancer. The majority of ovarian cancers are derived from epithelial cells, and 70% of patients with epithelial ovarian cancers present with late-stage disease. As a result, the long-term survival rate for individuals with this disease is very low. Identification of early-stage markers for ovarian cancer would significantly increase the survival rate. The molecular events that lead to ovarian cancer are poorly understood. Some of the known aberrations include mutation of p53 and microsatellite instability. Since gene expression patterns likely vary when normal ovary is compared to ovarian tumors, examination of gene expression in these tissues can identify possible markers for ovarian cancer.

Steroids are a class of lipid-soluble molecules, including cholesterol, bile acids, vitamin D, and hormones, that share a common four-ring structure based on cyclopentanoperhydrophenanthrene and that carry out a wide variety of functions. Cholesterol, for example, is a component of cell membranes that controls membrane fluidity. It is also a precursor for bile acids which solubilize lipids and facilitate absorption in the small intestine during digestion. Vitamin D regulates the absorption of calcium in the small intestine and controls the concentration of calcium in plasma. Steroid hormones, produced by the adrenal cortex, ovaries, and testes, include glucocorticoids, mineralocorticoids, androgens, and estrogens. They control various biological processes by binding to intracellular receptors that regulate transcription of specific genes in the nucleus. Glucocorticoids, for example, increase blood glucose concentrations by regulation of gluconeogenesis in the liver, increase blood concentrations of fatty acids by promoting lipolysis in adipose tissues, modulate

sensitivity to catcholamines in the central nervous system, and reduce inflammation. The principal mineralocorticoid, aldosterone, is produced by the adrenal cortex and acts on cells of the distal tubules of the kidney to enhance sodium ion reabsorption. Androgens, produced by the interstitial cells of Leydig in the testis, include the male sex hormone testosterone, which triggers changes at puberty, the production of sperm and maintenance of secondary sexual characteristics. Female sex hormones, estrogen and progesterone, are produced by the ovaries and also by the placenta and adrenal cortex of the fetus during pregnancy. Estrogen regulates female reproductive processes and secondary sexual characteristics. Progesterone regulates changes in the endometrium during the menstrual cycle and pregnancy.

Steroid hormones are widely used for fertility control and in anti-inflammatory treatments for physical injuries and diseases such as arthritis, asthma, and auto-immune disorders. Progesterone, a naturally occurring progestin, is primarily used to treat amenorrhea, abnormal uterine bleeding, or as a contraceptive. Endogenous progesterone is responsible for inducing secretory activity in the endometrium of the estrogen-primed uterus in preparation for the implantation of a fertilized egg and for the maintenance of pregnancy. It is secreted from the corpus luteum in response to luteinizing hormone (LH). The primary contraceptive effect of exogenous progestins involves the suppression of the midcycle surge of LH. At the cellular level, progestins diffuse freely into target cells and bind to the progesterone receptor. Target cells include the female reproductive tract, the mammary gland, the hypothalamus, and the pituitary. Once bound to the receptor, progestins slow the frequency of release of gonadotropin releasing hormone from the hypothalamus and blunt the pre-ovulatory LH surge, thereby preventing follicular maturation and ovulation. Progesterone has minimal estrogenic and androgenic activity. Progesterone is metabolized hepatically to pregnanediol and conjugated with glucuronic acid.

Medroxyprogesterone (MAH), also known as 6α-methyl-17-hydroxyprogesterone, is a synthetic progestin with a pharmacological activity about 15 times greater than progesterone. MAH is used for the treatment of renal and endometrial carcinomas, amenorrhea, abnormal uterine bleeding, and endometriosis associated with hormonal imbalance. MAH has a stimulatory effect on respiratory centers and has been used in cases of low blood oxygenation caused by sleep apnea, chronic obstructive pulmonary disease, or hypercapnia.

Danazol is a synthetic steroid derived from ethinyl testosterone. Danazol indirectly reduces estrogen production by lowering pituitary synthesis of follicle-stimulating hormone and LH. Danazol also binds to sex hormone receptors in target tissues, thereby exhibiting anabolic, antiestrognic, and weakly androgenic activity. Danazol does not possess any progestogenic activity, and does not suppress normal pituitary release of corticotropin or release of cortisol by the adrenal glands. Danazol is used in the treatment of endometriosis to relieve pain and inhibit endometrial cell growth.

It is also used to treat fibrocystic breast disease and hereditary angioedema.

Corticosteroids are used to relieve inflammation and to suppress the immune response. They inhibit eosinophil, basophil, and airway epithelial cell function by regulation of cytokines that mediate the inflammatory response. They inhibit leukocyte infiltration at the site of inflammation, interfere in the function of mediators of the inflammatory response, and suppress the humoral immune response. Corticosteroids are used to treat allergies, asthma, arthritis, and skin conditions. Beclomethasone is a synthetic glucocorticoid that is used to treat steroid-dependent asthma, to relieve symptoms associated with allergic or nonallergic (vasomotor) rhinitis, or to prevent recurrent nasal polyps following surgical removal. The anti-inflammatory and vasoconstrictive effects of intranasal beclomethasone are 5000 times greater than those produced by hydrocortisone.

The anti-inflammatory actions of corticosteroids are thought to involve phospholipase A_2 inhibitory proteins, collectively called lipocortins. Lipocortins, in turn, control the biosynthesis of potent mediators of inflammation such as prostaglandins and leukotrienes by inhibiting the release of the precursor molecule arachidonic acid. Proposed mechanisms of action include decreased IgE synthesis, increased number of β-adrenergic receptors on leukocytes, and decreased arachidonic acid metabolism. During an immediate allergic reaction, such as in chronic bronchial asthma, allergens bridge the IgE antibodies on the surface of mast cells, which triggers these cells to release chemotactic substances. Mast cell influx and activation, therefore, is partially responsible for the inflammation and hyperirritability of the oral mucosa in asthmatic patients. This inflammation can be retarded by administration of corticosteroids. ENDFIELD

The potential application of gene expression profiling is particularly relevant to measuring the toxic response to potential therapeutic compounds and the metabolic response to therapeutic agents. Diseases treated with steroids and disorders caused by the metabolic response to treatment with steroids include adenomatosis, cholestasis, cirrhosis, hemangioma, Henoch-Schonlein purpura, hepatitis, hepatocellular and metastatic carcinomas, idiopathic thrombocytopenic purpura, porphyria, sarcoidosis, and Wilson disease. Response may be measured by comparing both the levels and sequences expressed in tissues from subjects exposed to or treated with steroid compounds such as medroxyprogesterone (MAH) or budesonide (bude) with the levels and sequences expressed in normal untreated tissue.

The effects upon liver metabolism and hormone clearance mechanisms are important to understand the pharmacodynamics of a drug. The human C3A cell line is a clonal derivative of HepG2/C3 (hepatoma cell line, isolated from a 15-year-old male with liver tumor), which was selected for strong contact inhibition of growth. The use of a clonal population enhances the reproducibility of the cells. C3A cells have many characteristics of primary human hepatocytes in culture: i) expression of insulin receptor and insulin-like growth factor II receptor; ii) secretion of a

high ratio of serum albumin compared with α-fetoprotein iii) conversion of ammonia to urea and glutamine; iv) metabolism of aromatic amino acids; and v) proliferation in glucose-free and insulin-free medium. The C3A cell line is now well established as an in vitro model of the mature human liver (Mickelson et al. (1995) Hepatology 22:866-875; Nagendra et al. (1997) Am J Physiol 272:G408-G416).

There is a need in the art for new compositions, including nucleic acids and proteins, for the diagnosis, prevention, and treatment of cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders, diseases treated with steroids and disorders caused by the metabolic response to treatment with steroids.

SUMMARY OF THE INVENTION

Various embodiments of the invention provide purified polypeptides, molecules for disease detection and treatment, referred to collectively as "MDDT" and individually as "MDDT-1," "MDDT-2," "MDDT-3," "MDDT-4," "MDDT-5," "MDDT-6," "MDDT-7," "MDDT-8," "MDDT-9," "MDDT-10," "MDDT-11," "MDDT-12," "MDDT-13," "MDDT-14," "MDDT-15," "MDDT-16," "MDDT-17," "MDDT-18," "MDDT-19," "MDDT-20," "MDDT-21," "MDDT-22," "MDDT-23," "MDDT-24," "MDDT-25," "MDDT-26," "MDDT-27," "MDDT-28," "MDDT-29," "MDDT-30," "MDDT-31," "MDDT-32," "MDDT-33," "MDDT-34," "MDDT-35," "MDDT-36," "MDDT-37," "MDDT-38," "MDDT-39," "MDDT-40," "MDDT-41," "MDDT-42," "MDDT-43," "MDDT-44," "MDDT-45," "MDDT-46," "MDDT-47," "MDDT-48," "MDDT-49," "MDDT-50," "MDDT-51," "MDDT-52," "MDDT-53," "MDDT-54," "MDDT-55," and "MDDT-56" and methods for using these proteins and their encoding polynucleotides for the detection, diagnosis, and treatment of diseases and medical conditions. Embodiments also provide methods for utilizing the purified molecules for disease detection and treatment and/or their encoding polynucleotides for facilitating the drug discovery process, including determination of efficacy, dosage, toxicity, and pharmacology. Related embodiments provide methods for utilizing the purified molecules for disease detection and treatment and/or their encoding polynucleotides for investigating the pathogenesis of diseases and medical conditions.

An embodiment provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56. Another

embodiment provides an isolated polypeptide comprising an amino acid sequence of SEQ ID NO:1-56.

Still another embodiment provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56. In another embodiment, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-56. In an alternative embodiment, the polynucleotide is selected from the group consisting of SEQ ID NO:57-112.

Still another embodiment provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56. Another embodiment provides a cell transformed with the recombinant polynucleotide. Yet another embodiment provides a transgenic organism comprising the recombinant polynucleotide.

Another embodiment provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Yet another embodiment provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid

sequence selected from the group consisting of SEQ ID NO:1-56, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56.

Still yet another embodiment provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In other embodiments, the polynucleotide can comprise at least about 20, 30, 40, 60, 80, or 100 contiguous nucleotides.

Yet another embodiment provides a method for detecting a target polynucleotide in a sample, said target polynucleotide being selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex. In a related embodiment, the method can include detecting the amount of the hybridization complex. In still other embodiments, the probe can comprise at least about 20, 30, 40, 60, 80, or 100 contiguous nucleotides.

Still yet another embodiment provides a method for detecting a target polynucleotide in a sample, said target polynucleotide being selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method

comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof. In a related embodiment, the method can include detecting the amount of the amplified target polynucleotide or fragment thereof.

Another embodiment provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, and a pharmaceutically acceptable excipient. In one embodiment, the composition can comprise an amino acid sequence selected from the group consisting of SEQ ID NO:1-56. Other embodiments provide a method of treating a disease or condition associated with decreased or abnormal expression of functional MDDT, comprising administering to a patient in need of such treatment the composition.

Yet another embodiment provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. Another embodiment provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. Yet another embodiment provides a method of treating a disease or condition associated with decreased expression of functional MDDT, comprising administering to a patient in need of such treatment the composition.

Still yet another embodiment provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group

consisting of SEQ ID NO:1-56, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. Another embodiment provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. Yet another embodiment provides a method of treating a disease or condition associated with overexpression of functional MDDT, comprising administering to a patient in need of such treatment the composition.

Another embodiment provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

Yet another embodiment provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

Still yet another embodiment provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112, the method

comprising a) exposing a sample comprising the target polynucleotide to a compound, b) detecting altered expression of the target polynucleotide, and c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

Another embodiment provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide can comprise a fragment of a polynucleotide selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for full length polynucleotide and polypeptide embodiments of the invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptide embodiments of the invention. The probability scores for the matches between each polypeptide and its homolog(s) are also shown.

Table 3 shows structural features of polypeptide embodiments, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide embodiments, along with selected fragments of the polynucleotides.

Table 5 shows representative cDNA libraries for polynucleotide embodiments.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze polynucleotides and polypeptides, along with applicable descriptions, references, and threshold parameters.

Table 8 shows single nucleotide polymorphisms found in polynucleotide embodiments, along with allele frequencies in different human populations.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleic acids, and methods are described, it is understood that embodiments of the invention are not limited to the particular machines, instruments, materials, and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the invention.

As used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with various embodiments of the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"MDDT" refers to the amino acid sequences of substantially purified MDDT obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of MDDT. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other

compound or composition which modulates the activity of MDDT either by directly interacting with MDDT or by acting on components of the biological pathway in which MDDT participates.

An "allelic variant" is an alternative form of the gene encoding MDDT. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding MDDT include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as MDDT or a polypeptide with at least one functional characteristic of MDDT. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding MDDT, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide encoding MDDT. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent MDDT. Deliberate amino acid substitutions may be made on the basis of one or more similarities in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of MDDT is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" can refer to an oligopeptide, a peptide, a polypeptide, or a protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid. Amplification may be carried out using polymerase chain reaction (PCR) technologies or other nucleic acid amplification technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity

of MDDT. Antagonists may include proteins such as antibodies, anticalins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of MDDT either by directly interacting with MDDT or by acting on components of the biological pathway in which MDDT participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind MDDT polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an *in vitro* evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH_2), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker (Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13).

The term "intramer" refers to an aptamer which is expressed *in vivo*. For example, a vaccinia virus-based RNA expression system has been used to express specific RNA aptamers at high levels in the cytoplasm of leukocytes (Blind, M. et al. (1999) Proc. Natl. Acad. Sci. USA 96:3606-3610).

The term "spiegelmer" refers to an aptamer which includes L-DNA, L-RNA, or other left-

handed nucleotide derivatives or nucleotide-like molecules. Aptamers containing left-handed nucleotides are resistant to degradation by naturally occurring enzymes, which normally act on substrates containing right-handed nucleotides.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a polynucleotide having a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic MDDT, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide" and a "composition comprising a given polypeptide" can refer to any composition containing the given polynucleotide or polypeptide. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotides encoding MDDT or fragments of MDDT may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison

WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

"Exon shuffling" refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A "fragment" is a unique portion of MDDT or a polynucleotide encoding MDDT which can be identical in sequence to, but shorter in length than, the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from about 5 to about 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:57-112 can comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:57-112, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:57-112 can be employed in one or more embodiments of methods of the invention, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:57-112 from related polynucleotides. The precise length of a fragment of SEQ ID NO:57-112 and the region of SEQ ID NO:57-112 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-56 is encoded by a fragment of SEQ ID NO:57-112. A fragment of SEQ ID NO:1-56 can comprise a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-56. For example, a fragment of SEQ ID NO:1-56 can be used as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-56. The precise length of a fragment of SEQ ID NO:1-56 and the region of SEQ ID NO:1-56 to which the fragment corresponds can be determined based on the intended purpose for the fragment using one or more analytical methods described herein or otherwise known in the art.

A "full length" polynucleotide is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using one or more computer algorithms or programs known in the art or described herein. For example, percent identity can be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989; CABIOS 5:151-153) and in Higgins, D.G. et al. (1992; CABIOS 8:189-191). For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms which can be used is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 μg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of

the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 μg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acids by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid present in solution and another nucleic acid immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or polynucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of MDDT which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of MDDT which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, antibodies, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, antibody, or

other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of MDDT. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of MDDT.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an MDDT may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of MDDT.

"Probe" refers to nucleic acids encoding MDDT, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acids. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the

specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989; Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY), Ausubel, F.M. et al. (1999) Short Protocols in Molecular Biology, 4th ed., John Wiley & Sons, New York NY), and Innis, M. et al. (1990; PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA). PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a nucleic acid that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, *supra*. The term recombinant includes

nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA molecule, is composed of the same linear sequence of nucleotides as the reference DNA molecule with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing MDDT, nucleic acids encoding MDDT, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably at least about 75% free, and most preferably at least about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides

by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" or "expression profile" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. In another embodiment, the nucleic acid can be introduced by infection with a recombinant viral vector, such as a lentiviral vector (Lois, C. et al. (2002) Science 295:868-872). The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), *supra.*

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least

93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotides that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

Various embodiments of the invention include new human molecules for disease detection and treatment (MDDT), the polynucleotides encoding MDDT, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders, diseases treated with steroids and disorders caused by the metabolic response to treatment with steroids.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide embodiments of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown. Column 6 shows the Incyte ID numbers of physical, full length clones corresponding to

polypeptide and polynucleotide embodiments. The full length clones encode polypeptides which have at least 95% sequence identity to the polypeptides shown in column 3.

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (GenBank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability scores for the matches between each polypeptide and its homolog(s). Column 5 shows the annotation of the GenBank homolog(s) along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are molecules for disease detection and treatment. For example, SEQ ID NO:2 contains a potassium channel tetramerisation domain domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.)

In another example, SEQ ID NO:22 is 93% identical, from residue M1 to residue V1451, to mouse pecanex 1, which is the mouse homolog of Drosophila pecanex, a maternal-effect neurogenic protein (GenBank ID g6650377) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. Data from further BLAST analyses provide corroborative evidence that SEQ ID NO:22 is a pecanex 1 protein.

In another example, SEQ ID NO:31 is 33% identical, from residue R17 to residue G452, to Drosophila melanogaster Diablo (GenBank ID g7243777) as determined by the Basic Local Alignment Search- Tool (BLAST). (See Table 2.) The BLAST probability score is 2.9e-50, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:31 also contains a BTB-POZ protein interaction domain as determined by searching for

statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS and additional BLAST analyses provide further corroborative evidence that SEQ ID NO:31 is an apoptosis-associated protein.

In another example, SEQ ID NO:36 is 62% identical, from residue E84 to residue L370, to a human EVI-5 protein (GenBank ID g3093476) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 6.9e-90, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. Data from MOTIFS analysis provides further corroborative evidence that SEQ ID NO:36 is a protein with potential utility for disease detection or treatment.

In another example, SEQ ID NO:44 is 78% identical, from residue D224 to residue V838, and 98% identical, from residue M1 to residue W333, to human sporulation-induced transcript 4 (SIT4)-associated protein SAPLa (GenBank ID g11527201) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 6.7e-250 for the first homologous section and 1.8e-171 for the second, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. Data from other BLAST analyses provide further corroborative evidence that SEQ ID NO:44 is a cycle cell phosphorylation sit4-associating protein (a protein which associates with the sit4 phosphatase in a cell cycle-dependent manner).

In another example, SEQ ID NO:47 is 52% identical, from residue F6 to residue L256, to a WD-40-containing Xenopus laevis protein that is upregulated by thyroid hormone (GenBank ID g1314316) as determined by the Basic Local Alignment Search Tool (BLAST). The BLAST probability score is 6.3e-73, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:47 also contains a WD, G-beta repeat domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from MOTIFS and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:47 is a full-length human molecule for disease detection and treatment. SEQ ID NO:1, SEQ ID NO:3-21, SEQ ID NO:23-30, SEQ ID NO:32-35, SEQ ID NO:37-43, SEQ ID NO:45-46 and SEQ ID NO:48-56 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-56 are described in Table 7.

As shown in Table 4, the full length polynucleotide embodiments were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Column 1 lists the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:), the corresponding Incyte polynucleotide consensus sequence number (Incyte ID) for each polynucleotide of the invention, and the length of each polynucleotide sequence

in basepairs. Column 2 shows the nucleotide start (5′) and stop (3′) positions of the cDNA and/or genomic sequences used to assemble the full length polynucleotide embodiments, and of fragments of the polynucleotides which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:57-112 or that distinguish between SEQ ID NO:57-112 and related polynucleotides.

The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotides. In addition, the polynucleotide fragments described in column 2 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the polynucleotide fragments described in column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, a polynucleotide sequence identified as FL_$XXXXXX_N_1_N_2_YYYYY_N_3_N_4$ represents a "stitched" sequence in which *XXXXXX* is the identification number of the cluster of sequences to which the algorithm was applied, and *YYYYY* is the number of the prediction generated by the algorithm, and $N_{1,2,3\ldots}$, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, a polynucleotide sequence identified as FL*XXXXXX*_g*AAAAA*_g*BBBBB*_1_*N* is a "stretched" sequence, with *XXXXXX* being the Incyte project identification number, g*AAAAA* being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, g*BBBBB* being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and *N* referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, g*BBBBB*).

Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis methods associated with the prefixes (see Example IV and Example V).

Prefix	Type f analysis and/or examples of programs
GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).
INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in Table 4 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotides which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotides. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

Table 8 shows single nucleotide polymorphisms (SNPs) found in polynucleotide embodiments, along with allele frequencies in different human populations. Columns 1 and 2 show the polynucleotide sequence identification number (SEQ ID NO:) and the corresponding Incyte project identification number (PID) for polynucleotides of the invention. Column 3 shows the Incyte identification number for the EST in which the SNP was detected (EST ID), and column 4 shows the identification number for the SNP (SNP ID). Column 5 shows the position within the EST sequence at which the SNP is located (EST SNP), and column 6 shows the position of the SNP within the full-length polynucleotide sequence (CB1 SNP). Column 7 shows the allele found in the EST sequence. Columns 8 and 9 show the two alleles found at the SNP site. Column 10 shows the amino acid encoded by the codon including the SNP site, based upon the allele found in the EST. Columns 11-14 show the frequency of allele 1 in four different human populations. An entry of n/d (not detected) indicates that the frequency of allele 1 in the population was too low to be detected, while n/a (not available) indicates that the allele frequency was not determined for the population.

The invention also encompasses MDDT variants. A preferred MDDT variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the MDDT amino acid sequence, and which contains at least one functional or

structural characteristic of MDDT.

Various embodiments also encompass polynucleotides which encode MDDT. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:57-112, which encodes MDDT. The polynucleotide sequences of SEQ ID NO:57-112, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses variants of a polynucleotide encoding MDDT. In particular, such a variant polynucleotide will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a polynucleotide encoding MDDT. A particular aspect of the invention encompasses a variant of a polynucleotide comprising a sequence selected from the group consisting of SEQ ID NO:57-112 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:57-112. Any one of the polynucleotide variants described above can encode a polypeptide which contains at least one functional or structural characteristic of MDDT.

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide encoding MDDT. A splice variant may have portions which have significant sequence identity to a polynucleotide encoding MDDT, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing of exons during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to a polynucleotide encoding MDDT over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide encoding MDDT. For example, a polynucleotide comprising a sequence of SEQ ID NO:112 and a polynucleotide comprising a sequence of SEQ ID NO:59 are splice variants of each other. Any one of the splice variants described above can encode a polypeptide which contains at least one functional or structural characteristic of MDDT.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding MDDT, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the

polynucleotide sequence of naturally occurring MDDT, and all such variations are to be considered as being specifically disclosed.

Although polynucleotides which encode MDDT and its variants are generally capable of hybridizing to polynucleotides encoding naturally occurring MDDT under appropriately selected conditions of stringency, it may be advantageous to produce polynucleotides encoding MDDT or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding MDDT and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of polynucleotides which encode MDDT and MDDT derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic polynucleotide may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a polynucleotide encoding MDDT or any fragment thereof.

Embodiments of the invention can also include polynucleotides that are capable of hybridizing to the claimed polynucleotides, and, in particular, to those having the sequences shown in SEQ ID NO:57-112 and fragments thereof, under various conditions of stringency (Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511). Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Biosciences, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Invitrogen, Carlsbad CA). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Amersham Biosciences), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art (Ausubel et al., *supra*, ch. 7; Meyers, R.A. (1995) Molecular Biology and

Biotechnology, Wiley VCH, New York NY, pp. 856-853).

The nucleic acids encoding MDDT may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences (Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186). A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA (Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119). In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art (Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotides or fragments thereof which encode MDDT may be cloned in recombinant DNA molecules that direct expression of MDDT, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other polynucleotides which encode substantially the same or a functionally equivalent polypeptides may be produced and used to express MDDT.

The polynucleotides of the invention can be engineered using methods generally known in the art in order to alter MDDT-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of MDDT, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, polynucleotides encoding MDDT may be synthesized, in whole or in part, using one or more chemical methods well known in the art (Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232). Alternatively, MDDT itself or a fragment thereof may be synthesized using chemical methods known in the art. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques (Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; Roberge, J.Y. et al. (1995) Science 269:202-204). Automated

synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of MDDT, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography (Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421). The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (Creighton, *supra*, pp. 28-53).

In order to express a biologically active MDDT, the polynucleotides encoding MDDT or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotides encoding MDDT. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of polynucleotides encoding MDDT. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where a polynucleotide sequence encoding MDDT and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used (Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162).

Methods which are well known to those skilled in the art may be used to construct expression vectors containing polynucleotides encoding MDDT and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination (Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel et al., *supra*, ch. 1, 3, and 15).

A variety of expression vector/host systems may be utilized to contain and express polynucleotides encoding MDDT. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g.,

cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems (Sambrook, *supra*; Ausubel et al., *supra*; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355). Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of polynucleotides to the targeted organ, tissue, or cell population (Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5:350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90:6340-6344; Buller, R.M. et al. (1985) Nature 317:813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31:219-226; Verma, I.M. and N. Somia (1997) Nature 389:239-242). The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotides encoding MDDT. For example, routine cloning, subcloning, and propagation of polynucleotides encoding MDDT can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Invitrogen). Ligation of polynucleotides encoding MDDT into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence (Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509). When large quantities of MDDT are needed, e.g. for the production of antibodies, vectors which direct high level expression of MDDT may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of MDDT. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign polynucleotide sequences into the host genome for stable propagation (Ausubel et al., *supra*; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; Scorer, C.A. et al. (1994) Bio/Technology 12:181-184).

Plant systems may also be used for expression of MDDT. Transcription of polynucleotides encoding MDDT may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J.

6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection (The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196).

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, polynucleotides encoding MDDT may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses MDDT in host cells (Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes (Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355).

For long term production of recombinant proteins in mammalian systems, stable expression of MDDT in cell lines is preferred. For example, polynucleotides encoding MDDT can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ and *apr*⁻ cells, respectively (Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823). Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol.

150:1-14). Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites (Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051). Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β-glucuronidase and its substrate β-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding MDDT is inserted within a marker gene sequence, transformed cells containing polynucleotides encoding MDDT can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding MDDT under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the polynucleotide encoding MDDT and that express MDDT may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of MDDT using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on MDDT is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art (Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding MDDT include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, polynucleotides encoding MDDT, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially

available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Biosciences, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with polynucleotides encoding MDDT may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode MDDT may be designed to contain signal sequences which direct secretion of MDDT through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted polynucleotides or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant polynucleotides encoding MDDT may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric MDDT protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of MDDT activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the MDDT encoding sequence and the heterologous protein

sequence, so that MDDT may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel et al. (*supra*, ch. 10 and 16). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In another embodiment, synthesis of radiolabeled MDDT may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

MDDT, fragments of MDDT, or variants of MDDT may be used to screen for compounds that specifically bind to MDDT. One or more test compounds may be screened for specific binding to MDDT. In various embodiments, 1, 2, 3, 4, 5, 10, 20, 50, 100, or 200 test compounds can be screened for specific binding to MDDT. Examples of test compounds can include antibodies, anticalins, oligonucleotides, proteins (e.g., ligands or receptors), or small molecules.

In related embodiments, variants of MDDT can be used to screen for binding of test compounds, such as antibodies, to MDDT, a variant of MDDT, or a combination of MDDT and/or one or more variants MDDT. In an embodiment, a variant of MDDT can be used to screen for compounds that bind to a variant of MDDT, but not to MDDT having the exact sequence of a sequence of SEQ ID NO:1-56. MDDT variants used to perform such screening can have a range of about 50% to about 99% sequence identity to MDDT, with various embodiments having 60%, 70%, 75%, 80%, 85%, 90%, and 95% sequence identity.

In an embodiment, a compound identified in a screen for specific binding to MDDT can be closely related to the natural ligand of MDDT, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner (Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2):Chapter 5). In another embodiment, the compound thus identified can be a natural ligand of a receptor MDDT (Howard, A.D. et al. (2001) Trends Pharmacol. Sci.22:132-140; Wise, A. et al. (2002) Drug Discovery Today 7:235-246).

In other embodiments, a compound identified in a screen for specific binding to MDDT can be closely related to the natural receptor to which MDDT binds, at least a fragment of the receptor, or a fragment of the receptor including all or a portion of the ligand binding site or binding pocket. For example, the compound may be a receptor for MDDT which is capable of propagating a signal, or a decoy receptor for MDDT which is not capable of propagating a signal (Ashkenazi, A. and V.M. Divit (1999) Curr. Opin. Cell Biol. 11:255-260; Mantovani, A. et al. (2001) Trends Immunol. 22:328-336). The compound can be rationally designed using known techniques. Examples of such techniques include those used to construct the compound etanercept (ENBREL; Immunex Corp.,

Seattle WA), which is efficacious for treating rheumatoid arthritis in humans. Etanercept is an engineered p75 tumor necrosis factor (TNF) receptor dimer linked to the Fc portion of human IgG_1 (Taylor, P.C. et al. (2001) Curr. Opin. Immunol. 13:611-616).

In one embodiment, two or more antibodies having similar or, alternatively, different specificities can be screened for specific binding to MDDT, fragments of MDDT, or variants of MDDT. The binding specificity of the antibodies thus screened can thereby be selected to identify particular fragments or variants of MDDT. In one embodiment, an antibody can be selected such that its binding specificity allows for preferential identification of specific fragments or variants of MDDT. In another embodiment, an antibody can be selected such that its binding specificity allows for preferential diagnosis of a specific disease or condition having increased, decreased, or otherwise abnormal production of MDDT.

In an embodiment, anticalins can be screened for specific binding to MDDT, fragments of MDDT, or variants of MDDT. Anticalins are ligand-binding proteins that have been constructed based on a lipocalin scaffold (Weiss, G.A. and H.B. Lowman (2000) Chem. Biol. 7:R177-R184; Skerra, A. (2001) J. Biotechnol. 74:257-275). The protein architecture of lipocalins can include a beta-barrel having eight antiparallel beta-strands, which supports four loops at its open end. These loops form the natural ligand-binding site of the lipocalins, a site which can be re-engineered *in vitro* by amino acid substitutions to impart novel binding specificities. The amino acid substitutions can be made using methods known in the art or described herein, and can include conservative substitutions (e.g., substitutions that do not alter binding specificity) or substitutions that modestly, moderately, or significantly alter binding specificity.

In one embodiment, screening for compounds which specifically bind to, stimulate, or inhibit MDDT involves producing appropriate cells which express MDDT, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing MDDT or cell membrane fractions which contain MDDT are then contacted with a test compound and binding, stimulation, or inhibition of activity of either MDDT or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with MDDT, either in solution or affixed to a solid support, and detecting the binding of MDDT to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

An assay can be used to assess the ability of a compound to bind to its natural ligand and/or to inhibit the binding of its natural ligand to its natural receptors. Examples of such assays include radio-labeling assays such as those described in U.S. Patent No. 5,914,236 and U.S. Patent No. 6,372,724. In a related embodiment, one or more amino acid substitutions can be introduced into a polypeptide compound (such as a receptor) to improve or alter its ability to bind to its natural ligands (Matthews, D.J. and J.A. Wells. (1994) Chem. Biol. 1:25-30). In another related embodiment, one or more amino acid substitutions can be introduced into a polypeptide compound (such as a ligand) to improve or alter its ability to bind to its natural receptors (Cunningham, B.C. and J.A. Wells (1991) Proc. Natl. Acad. Sci. USA 88:3407-3411; Lowman, H.B. et al. (1991) J. Biol. Chem. 266:10982-10988).

MDDT, fragments of MDDT, or variants of MDDT may be used to screen for compounds that modulate the activity of MDDT. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for MDDT activity, wherein MDDT is combined with at least one test compound, and the activity of MDDT in the presence of a test compound is compared with the activity of MDDT in the absence of the test compound. A change in the activity of MDDT in the presence of the test compound is indicative of a compound that modulates the activity of MDDT. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising MDDT under conditions suitable for MDDT activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of MDDT may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding MDDT or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease (see, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337). For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (*neo*; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce

heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding MDDT may also be manipulated *in vitro* in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding MDDT can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding MDDT is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress MDDT, e.g., by secreting MDDT in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of MDDT and molecules for disease detection and treatment. In particular, SEQ ID NO:110 shows co-expression with osteoporosis-relevant genes. In addition, examples of tissues expressing MDDT can be found in Table 6 and can also be found in Example XI. Therefore, MDDT appears to play a role in cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders, diseases treated with steroids and disorders caused by the metabolic response to treatment with steroids. In the treatment of disorders associated with increased MDDT expression or activity, it is desirable to decrease the expression or activity of MDDT. In the treatment of disorders associated with decreased MDDT expression or activity, it is desirable to increase the expression or activity of MDDT.

Therefore, in one embodiment, MDDT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MDDT. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate,

salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a disease treated with a steroid and a disorder caused by the metabolic response to treatment with steroids, such as adenomatosis, cholestasis, cirrhosis, hemangioma, Henoch-Schonlein purpura, hepatitis, hepatocellular and metastatic carcinomas, idiopathic thrombocytopenic purpura, porphyria, sarcoidosis, and Wilson disease; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal

disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia. SEQ ID NO:2 can be used in the diagnosis and treatment of Tangier disease and SEQ ID NO:5 can be used in the diagnosis and treatment of type II diabetes.

In another embodiment, a vector capable of expressing MDDT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MDDT including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified MDDT in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MDDT including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of MDDT may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MDDT including, but not limited to, those listed above.

In a further embodiment, an antagonist of MDDT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of MDDT. Examples of such disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders, diseases treated with steroids and disorders caused by the metabolic response to treatment with steroids described above. In one aspect, an antibody which specifically binds MDDT may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express MDDT.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding MDDT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of MDDT including, but not limited to, those described above.

In other embodiments, any protein, agonist, antagonist, antibody, complementary sequence, or vector embodiments may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with

lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of MDDT may be produced using methods which are generally known in the art. In particular, purified MDDT may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind MDDT. Antibodies to MDDT may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use. Single chain antibodies (e.g., from camels or llamas) may be potent enzyme inhibitors and may have advantages in the design of peptide mimetics, and in the development of immuno-adsorbents and biosensors (Muyldermans, S. (2001) J. Biotechnol. 74:277-302).

For the production of antibodies, various hosts including goats, rabbits, rats, mice, camels, dromedaries, llamas, humans, and others may be immunized by injection with MDDT or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to MDDT have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of MDDT amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to MDDT may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; Takeda, S. et al. (1985)

Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce MDDT-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137).

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for MDDT may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W.D. et al. (1989) Science 246:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between MDDT and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering MDDT epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for MDDT. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of MDDT-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple MDDT epitopes, represents the average affinity, or avidity, of the antibodies for MDDT. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular MDDT epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the MDDT-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of MDDT, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical

Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of MDDT-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available (Catty, *supra*; Coligan et al., *supra*).

In another embodiment of the invention, polynucleotides encoding MDDT, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding MDDT. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding MDDT (Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press, Totawa NJ).

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein (Slater, J.E. et al. (1998) J. Allergy Clin. Immunol. 102:469-475; Scanlon, K.J. et al. (1995) 9:1288-1296). Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors (Miller, A.D. (1990) Blood 76:271; Ausubel et al., *supra*; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63:323-347). Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art (Rossi, J.J. (1995) Br. Med. Bull. 51:217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87:1308-1315; Morris, M.C. et al. (1997) Nucleic Acids Res. 25:2730-2736).

In another embodiment of the invention, polynucleotides encoding MDDT may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial

hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in MDDT expression or regulation causes disease, the expression of MDDT from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in MDDT are treated by constructing mammalian expression vectors encoding MDDT and introducing these vectors by mechanical means into MDDT-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J.-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of MDDT include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). MDDT may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding MDDT from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental

parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to MDDT expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding MDDT under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In an embodiment, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding MDDT to cells which have one or more genetic abnormalities with respect to the expression of MDDT. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999; Annu. Rev. Nutr. 19:511-544) and Verma, I.M. and N. Somia (1997; Nature 18:389:239-242).

In another embodiment, a herpes-based, gene therapy delivery system is used to deliver

polynucleotides encoding MDDT to target cells which have one or more genetic abnormalities with respect to the expression of MDDT. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing MDDT to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999; J. Virol. 73:519-532) and Xu, H. et al. (1994; Dev. Biol. 163:152-161). The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another embodiment, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding MDDT to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for MDDT into the alphavirus genome in place of the capsid-coding region results in the production of a large number of MDDT-coding RNAs and the synthesis of high levels of MDDT in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of MDDT into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the

art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177). A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of RNA molecules encoding MDDT.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA molecules encoding MDDT. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine,

cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding MDDT. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased MDDT expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding MDDT may be therapeutically useful, and in the treatment of disorders associated with decreased MDDT expression or activity, a compound which specifically promotes expression of the polynucleotide encoding MDDT may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding MDDT is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding MDDT are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding MDDT. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a

combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art (Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466).

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of MDDT, antibodies to MDDT, and mimetics, agonists, antagonists, or inhibitors of MDDT.

The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of

macromolecules comprising MDDT or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, MDDT or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example MDDT or fragments thereof, antibodies of MDDT, and agonists, antagonists or inhibitors of MDDT, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells,

conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind MDDT may be used for the diagnosis of disorders characterized by expression of MDDT, or in assays to monitor patients being treated with MDDT or agonists, antagonists, or inhibitors of MDDT. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for MDDT include methods which utilize the antibody and a label to detect MDDT in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring MDDT, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of MDDT expression. Normal or standard values for MDDT expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to MDDT under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of MDDT expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, polynucleotides encoding MDDT may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotides, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of MDDT may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of MDDT, and to monitor regulation of MDDT levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotides, including genomic sequences, encoding MDDT or closely related molecules may be used to identify nucleic acid sequences which encode MDDT. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding MDDT, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the MDDT encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:57-112 or from genomic sequences including promoters, enhancers, and introns of the MDDT gene.

Means for producing specific hybridization probes for polynucleotides encoding MDDT

include the cloning of polynucleotides encoding MDDT or MDDT derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotides encoding MDDT may be used for the diagnosis of disorders associated with expression of MDDT. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a disease treated with a steroid and a disorder caused by the metabolic response to treatment with steroids, such as adenomatosis, cholestasis, cirrhosis, hemangioma, Henoch-Schonlein purpura, hepatitis, hepatocellular and metastatic carcinomas, idiopathic thrombocytopenic purpura, porphyria, sarcoidosis, and Wilson disease; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial

dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia. SEQ ID NO:58, encoding SEQ ID NO:2, and SEQ ID NO:2 can be used in the diagnosis and treatment of Tangier disease and SEQ ID NO:61, encoding SEQ ID NO:5, and SEQ ID NO:5 can be used in the diagnosis and treatment of type II diabetes. Polynucleotides encoding MDDT may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered MDDT expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, polynucleotides encoding MDDT may be used in assays that detect the presence of associated disorders, particularly those mentioned above. Polynucleotides complementary to sequences encoding MDDT may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of polynucleotides encoding

MDDT in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of MDDT, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding MDDT, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier, thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding MDDT may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding MDDT, or a fragment of a polynucleotide complementary to the polynucleotide encoding MDDT, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from polynucleotides encoding MDDT may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from

polynucleotides encoding MDDT are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

SNPs may be used to study the genetic basis of human disease. For example, at least 16 common SNPs have been associated with non-insulin-dependent diabetes mellitus. SNPs are also useful for examining differences in disease outcomes in monogenic disorders, such as cystic fibrosis, sickle cell anemia, or chronic granulomatous disease. For example, variants in the mannose-binding lectin, MBL2, have been shown to be correlated with deleterious pulmonary outcomes in cystic fibrosis. SNPs also have utility in pharmacogenomics, the identification of genetic variants that influence a patient's response to a drug, such as life-threatening toxicity. For example, a variation in N-acetyl transferase is associated with a high incidence of peripheral neuropathy in response to the anti-tuberculosis drug isoniazid, while a variation in the core promoter of the ALOX5 gene results in diminished clinical response to treatment with an anti-asthma drug that targets the 5-lipoxygenase pathway. Analysis of the distribution of SNPs in different populations is useful for investigating genetic drift, mutation, recombination, and selection, as well as for tracing the origins of populations and their migrations (Taylor, J.G. et al. (2001) Trends Mol. Med. 7:507-512; Kwok, P.-Y. and Z. Gu (1999) Mol. Med. Today 5:538-543; Nowotny, P. et al. (2001) Curr. Opin. Neurobiol. 11:637-641).

Methods which may also be used to quantify the expression of MDDT include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves (Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236). The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the

polynucleotides described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, MDDT, fragments of MDDT, or antibodies specific for MDDT may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time (Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484; hereby expressly incorporated by reference herein). Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression *in vivo*, as in the case of a tissue or biopsy sample, or *in vitro*, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with *in vitro* model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471). If a test compound has a signature similar to that of a

compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity (see, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at http://www.niehs.nih.gov/oc/news/toxchip.htm). Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In an embodiment, the toxicity of a test compound can be assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another embodiment relates to the use of the polypeptides disclosed herein to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially

sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of interest. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for MDDT to quantify the levels of MDDT expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art (Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662). Various types of microarrays are well known and thoroughly described in Schena, M., ed. (1999; DNA Microarrays: A Practical Approach, Oxford University Press, London).

In another embodiment of the invention, nucleic acid sequences encoding MDDT may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries (Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; Trask, B.J. (1991) Trends Genet. 7:149-154). Once mapped, the nucleic acid sequences may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP) (Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357).

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical and genetic map data (Heinz-Ulrich, et al. (1995) in Meyers, *supra*, pp. 965-968). Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding MDDT on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further

investigation (Gatti, R.A. et al. (1988) Nature 336:577-580). The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, MDDT, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between MDDT and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest (Geysen, et al. (1984) PCT application WO84/03564). In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with MDDT, or fragments thereof, and washed. Bound MDDT is then detected by methods well known in the art. Purified MDDT can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding MDDT specifically compete with a test compound for binding MDDT. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with MDDT.

In additional embodiments, the nucleotide sequences which encode MDDT may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, in particular U.S. Ser. No. 60/304,298, U.S. Ser. No. 60/305,324, U.S. Ser. No. 60/307,003, U.S. Ser. No. 60/308,185, U.S. Ser. No. 60/310,096, U.S. Ser. No. 60/311,551 and U.S. Ser. No. 60/363,649, are expressly incorporated by reference herein.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database

(Incyte Genomics, Palo Alto CA). Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Invitrogen), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Invitrogen), using the recommended procedures or similar methods known in the art (Ausubel et al., *supra*, ch. 5). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Biosciences) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Invitrogen), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Invitrogen.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Biosciences or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Amersham Biosciences); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (Ausubel et al., *supra*, ch. 7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM; PROTEOME databases with sequences from *Homo sapiens, Rattus norvegicus, Mus musculus, Caenorhabditis elegans, Saccharomyces cerevisiae, Schizosaccharomyces pombe*, and *Candida albicans* (Incyte Genomics, Palo Alto CA); hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM (Haft, D.H. et al. (2001) Nucleic Acids Res. 29:41-43); and HMM-based protein domain databases such as SMART (Schultz, J. et al. (1998) Proc. Natl. Acad. Sci. USA 95:5857-5864; Letunic, I. et al. (2002) Nucleic Acids Res. 30:242-244). (HMM is a probabilistic approach which analyzes consensus primary structures of gene families; see, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and

HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, the PROTEOME databases, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM; and HMM-based protein domain databases such as SMART. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:57-112. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative molecules for disease detection and treatment were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94; Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates

predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode molecules for disease detection and treatment, the encoded polypeptides were analyzed by querying against PFAM models for molecules for disease detection and treatment. Potential molecules for disease detection and treatment were also identified by homology to Incyte cDNA sequences that had been annotated as molecules for disease detection and treatment. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

<u>"Stitched" Sequences</u>

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent

type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

<u>**"Stretched" Sequences**</u>

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of MDDT Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:57-112 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:57-112 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site

(http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound (Sambrook, *supra*, ch. 7; Ausubel et al., *supra*, ch. 4).

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score x Percent Identity}}{\text{5 x minimum \{length(Seq. 1), length(Seq. 2)\}}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotides encoding MDDT are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male;

germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding MDDT. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of MDDT Encoding Polynucleotides

Full length polynucleotides are produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Biosciences), ELONGASE enzyme (Invitrogen), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the

concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Biosciences). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Biosciences), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Biosciences) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Biosciences) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotides are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Identification of Single Nucleotide Polymorphisms in MDDT Encoding Polynucleotides

Common DNA sequence variants known as single nucleotide polymorphisms (SNPs) were identified in SEQ ID NO:57-112 using the LIFESEQ database (Incyte Genomics). Sequences from the same gene were clustered together and assembled as described in Example III, allowing the identification of all sequence variants in the gene. An algorithm consisting of a series of filters was used to distinguish SNPs from other sequence variants. Preliminary filters removed the majority of basecall errors by requiring a minimum Phred quality score of 15, and removed sequence alignment errors and errors resulting from improper trimming of vector sequences, chimeras, and splice variants. An automated procedure of advanced chromosome analysis analysed the original chromatogram files in the vicinity of the putative SNP. Clone error filters used statistically generated

algorithms to identify errors introduced during laboratory processing, such as those caused by reverse transcriptase, polymerase, or somatic mutation. Clustering error filters used statistically generated algorithms to identify errors resulting from clustering of close homologs or pseudogenes, or due to contamination by non-human sequences. A final set of filters removed duplicates and SNPs found in immunoglobulins or T-cell receptors.

Certain SNPs were selected for further characterization by mass spectrometry using the high throughput MASSARRAY system (Sequenom, Inc.) to analyze allele frequencies at the SNP sites in four different human populations. The Caucasian population comprised 92 individuals (46 male, 46 female), including 83 from Utah, four French, three Venezualan, and two Amish individuals. The African population comprised 194 individuals (97 male, 97 female), all African Americans. The Hispanic population comprised 324 individuals (162 male, 162 female), all Mexican Hispanic. The Asian population comprised 126 individuals (64 male, 62 female) with a reported parental breakdown of 43% Chinese, 31% Japanese, 13% Korean, 5% Vietnamese, and 8% other Asian. Allele frequencies were first analyzed in the Caucasian population; in some cases those SNPs which showed no allelic variance in this population were not further tested in the other three populations.

X. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:57-112 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Biosciences), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Biosciences). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

XI. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing

photolithography, piezoelectric printing (ink-jet printing; see, e.g., Baldeschweiler et al., *supra*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena, M., ed. (1999) DNA Microarrays: A Practical Approach, Oxford University Press, London). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements (Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31).

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)$^+$ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)$^+$ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Biosciences). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)$^+$ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)$^+$ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated

using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Biosciences).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in U.S. Patent No. 5,807,522, incorporated herein by reference. 1 μl of the array element DNA, at an average concentration of 100 ng/μl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60° C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μl of sample mixture consisting of 0.2 μg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65° C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60° C. The arrays are washed for 10 min at 45° C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45° C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte). Array elements that exhibited at least about a two-fold change in expression, a signal-to-background ratio of at least 2.5, and an element spot size of at least 40% were identified as differentially

expressed using the GEMTOOLS program (Incyte Genomics).

Expression

For example, SEQ ID NO:58 was downregulated by at least two-fold in five out of the eight endothelial cell lines treated with 10ng/ml TNF-α within four hours. SEQ ID NO:58 was also down regulated in Tangier disease fibroblasts as compared to normal controls suggesting that SEQ ID NO:58, encoding SEQ ID NO:2, can be used for the diagnosis, prognosis or treatment of a cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, coronary artery bypass,congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and Tangier disease.

For example, SEQ ID NO:90 showed differential expression in treated versus non-treated C3A cells as determined by microarray analysis. The expression of SEQ ID NO:90 was decreased by at least two fold in C3A cells treated with 10 μM MAH for one to six hours and with 100 μM for one or six hours versus untreated C3A cells. SEQ ID NO:90 expression was also decreased by at least two-fold when C3A cells were treated for one to three hours with either 1μM, 10 μM, or 100 μM bude when compared with untreated C3A cells. These experiments indicate that SEQ ID NO:90 was significantly under-expressed in C3A cells when tested with two steroid compounds, further establishing the utility of SEQ ID NO:90 as a diagnostic marker or as a potential therapeutic target for liver disorders associated with steroid therapy such as adenomatosis, cholestasis, cirrhosis,

hemangioma, Henoch-Schonlein purpura, hepatitis, hepatocellular and metastatic carcinomas, idiopathic thrombocytopenic purpura, porphyria, sarcoidosis, and Wilson disease.

Human CD34 positive precursor cells were isolated by positive immunomagnetic selection from the leukapheresis of normal volunteer donors who had undergone G-CSF-induced stem cell mobilization. The purified CD34+ cells were cultured in vitro for 10 days in the presence of recombinant GM-CSF, Stem Cell Factor, TNF-alpha, TGF-beta1, and Flt3-Ligand. The resulting expanded cell population was enriched for cell cluster-forming immature dendritic cells (Lci) by sedimentation over a 7.5 % BSA column at 1 g for 30 min. Immature dendritic cells were cultured for two additional days in the presence of the same combination of cytokines supplemented with LPS, IL-1beta, TNF-alpha, or double strand RNA. In addition, cluster-forming immature dendritic cells were disrupted by vigorous pipetting and cultured for two additional days in the presence of the same combination of cytokines without addition of any additional factor. The partially mature dendritic cells derived by mechanical disruption of cell clusters are characterized by the presence of intracellular rod-shaped structures called Birbeck's Granules. The dendritic cell population produced using this method was called Birbeck's Granule-positive dendritic cells, or BG.

CD34+ precursor cells were compared to immature dendritic cells (Lci); Lci were compared to mature dendritic cells derived in the presence of LPS, IL-1b, TNF-alpha, or double strand RNA; and undisturbed Lci (Clusters) were compared to BG. Array elements that exhibited about at least a two-fold change in expression and a signal intensity over 250 units, a signal-to-background ratio of at least 2.5, and an element spot size of at least 40% were identified as differentially expressed using the GEMTOOLS program (Incyte Genomics). SEQ ID NO:96, which contains a GTPase activating protein motif for Arf and a Rho GAP domain, showed at least a two-fold increased expression during the differentiation of dendritic cells when induced by simple mechanical disaggregation and a greater than two-fold expression during maturation of these cells when induced by TNF-alpha. TNF-alpha is a factor produced by many cell types in response to stress. In addition, mechanical disruption is a significant stress factor during organ transplantation. Further, these experiments indicate that SEQ ID NO:96 was significantly over-expressed during differentiation and maturation of human dendritic cells, further establishing the utility of SEQ ID NO:96 as a diagnostic marker or as a potential therapeutic target for organ transplant disorders.

For example, SEQ ID NO:110 showed differential expression in diseased versus normal tissue as determined by microarray analysis. Matched normal ovary and ovarian tumor tissue samples are provided by the Huntsman Cancer Institute, (Salt Lake City, UT). The expression of MDDT was decreased in ovarian tumor cells relative to non-tumorous ovarian cells. Therefore, SEQ ID NO:110 is useful in diagnostic assays for ovarian cancer.

XII. Complementary Polynucleotides

Sequences complementary to the MDDT-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring MDDT. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of MDDT. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the MDDT-encoding transcript.

XIII. Expression of MDDT

Expression and purification of MDDT is achieved using bacterial or virus-based expression systems. For expression of MDDT in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express MDDT upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of MDDT in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding MDDT by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus (Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945).

In most expression systems, MDDT is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Biosciences). Following purification, the GST moiety can be proteolytically cleaved from MDDT at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-

His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel et al. (*supra*, ch. 10 and 16). Purified MDDT obtained by these methods can be used directly in the assays shown in Examples XVII, and XVIII where applicable.

XIV. Functional Assays

MDDT function is assessed by expressing the sequences encoding MDDT at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT plasmid (Invitrogen, Carlsbad CA) and PCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994; Flow Cytometry, Oxford, New York NY).

The influence of MDDT on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding MDDT and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding MDDT and other genes of interest can be analyzed by northern analysis or microarray techniques.

XV. Production of MDDT Specific Antibodies

MDDT substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize animals (e.g., rabbits, mice, etc.) and to produce antibodies using standard protocols.

Alternatively, the MDDT amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art (Ausubel et al., *supra*, ch. 11).

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity (Ausubel et al., *supra*). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-MDDT activity by, for example, binding the peptide or MDDT to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XVI. Purification of Naturally Occurring MDDT Using Specific Antibodies

Naturally occurring or recombinant MDDT is substantially purified by immunoaffinity chromatography using antibodies specific for MDDT. An immunoaffinity column is constructed by covalently coupling anti-MDDT antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Biosciences). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing MDDT are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of MDDT (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/MDDT binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and MDDT is collected.

XVII. Identification of Molecules Which Interact with MDDT

MDDT, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent (Bolton, A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled MDDT, washed, and any wells with labeled MDDT complex are assayed. Data obtained using different concentrations of MDDT are used to calculate values for the number, affinity, and association of MDDT with the candidate molecules.

Alternatively, molecules interacting with MDDT are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989; Nature 340:245-246), or using commercially

available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

MDDT may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVIII. Demonstration of MDDT Activity

Phorbol ester binding activity of MDDT is measured using an assay based on the fluorescent phorbol ester sapinotoxin-D (SAPD). Binding of SAPD to MDDT is quantified by measuring the resonance energy transfer from MDDT tryptophans to the 2-(N-methylamino)benzoyl fluorophore of the phorbol ester, as described by Slater et al. ((1996) J. Biol. Chem. 271:4627-4631).

Various modifications and variations of the described compositions, methods, and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. It will be appreciated that the invention provides novel and useful proteins, and their encoding polynucleotides, which can be used in the drug discovery process, as well as methods for using these compositions for the detection, diagnosis, and treatment of diseases and conditions. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Nor should the description of such embodiments be considered exhaustive or limit the invention to the precise forms disclosed. Furthermore, elements from one embodiment can be readily recombined with elements from one or more other embodiments. Such combinations can form a number of embodiments within the scope of the invention. It is intended that the scope of the invention be defined by the following claims and their equivalents.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Incyte Full Length Clones
2867236	1	2867236CD1	57	2867236CB1	1536179CA2
1294096	2	1294096CD1	58	1294096CB1	
7238537	3	7238537CD1	59	7238537CB1	90171762CA2, 90171778CA2, 90171786CA2, 90171870CA2, 90171886CA2, 90171894CA2, 90190048CA2, 90190116CA2, 90190124CA2, 90190132CA2
7494391	4	7494391CD1	60	7494391CB1	90140260CA2
6451054	5	6451054CD1	61	6451054CB1	
7494592	6	7494592CD1	62	7494592CB1	
5202657	7	5202657CD1	63	5202657CB1	3343965CA2
2013529	8	2013529CD1	64	2013529CB1	4010537CA2, 6300041CA2
3841351	9	3841351CD1	65	3841351CB1	
152116	10	152116CD1	66	152116CB1	
2381031	11	2381031CD1	67	2381031CB1	2381031CA2
2511371	12	2511371CD1	68	2511371CB1	
8068623	13	8068623CD1	69	8068623CB1	
677977	14	677977CD1	70	677977CB1	
1661472	15	1661472CD1	71	1661472CB1	
1748508	16	1748508CD1	72	1748508CB1	90132493CA2
2159545	17	2159545CD1	73	2159545CB1	
8560269	18	8560269CD1	74	8560269CB1	
8710302	19	8710302CD1	75	8710302CB1	
6778214	20	6778214CD1	76	6778214CB1	
258383	21	258383CD1	77	258383CB1	90140053CA2, 90140161CA2
2804937	22	2804937CD1	78	2804937CB1	
7494915	23	7494915CD1	79	7494915CB1	
2073751	24	2073751CD1	80	2073751CB1	
3178841	25	3178841CD1	81	3178841CB1	
3674807	26	3674807CD1	82	3674807CB1	3674807CA2
1794922	27	1794922CD1	83	1794922CB1	90144984CA2
1795509	28	1795509CD1	84	1795509CB1	690351CA2, 90131912CA2, 90131949CA2, 90131952CA2, 90131960CA2, 90131976CA2, 90131992CA2, 90132060CA2, 90132084CA2, 90132092CA2

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Incyte Full Length Clones
2017180	29	2017180CD1	85	2017180CB1	2807727CA2
219442	30	219442CD1	86	219442CB1	
2597459	31	2597459CD1	87	2597459CB1	90140160CA2
2783863	32	2783863CD1	88	2783863CB1	
2902971	33	2902971CD1	89	2902971CB1	
368660	34	368660CD1	90	368660CB1	90130001CA2
2804990	35	2804990CD1	91	2804990CB1	7616219CA2
168571	36	168571CD1	92	168571CB1	
1286391	37	1286391CD1	93	1286391CB1	
2007684	38	2007684CD1	94	2007684CB1	
2227040	39	2227040CD1	95	2227040CB1	
4346130	40	4346130CD1	96	4346130CB1	
55117040	41	55117040CD1	97	55117040CB1	55117036CA2
7472392	42	7472392CD1	98	7472392CB1	6622373CA2
4028960	43	4028960CD1	99	4028960CB1	
8227004	44	8227004CD1	100	8227004CB1	3279686CA2
3044763	45	3044763CD1	101	3044763CB1	90126287CA2
4044519	46	4044519CD1	102	4044519CB1	4044519CA2, 90106511CA2, 90106619CA2, 90106627CA2, 90106659CA2
71351918	47	71351918CD1	103	71351918CB1	
8109363	48	8109363CD1	104	8109363CB1	3853651CA2, 6859649CA2
1272746	49	1272746CD1	105	1272746CB1	
1839974	50	1839974CD1	106	1839974CB1	90120531CA2
1877336	51	1877336CD1	107	1877336CB1	
2321054	52	2321054CD1	108	2321054CB1	1236972CA2, 1398127CA2, 2245649CA2, 2321054CA2, 90106305CA2, 90106337CA2, 90106353CA2, 90106361CA2, 90106369CA2, 90106377CA2, 90106385CA2, 90106393CA2, 90106405CA2, 90106429CA2, 90106437CA2, 90106469CA2, 90106493CA2
2796034	53	2796034CD1	109	2796034CB1	
4413112	54	4413112CD1	110	4413112CB1	90108176CA2, 90172662CA2

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Incyte Full Length Clones
7654832	55	7654832CD1	111	7654832CB1	90110764CA2, 90110788CA2, 90110880CA2
7503849	56	7503849CD1	112	7503849CB1	

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
1	2867236CD1	g11065721	3.50E-134	[Homo sapiens] 28kD interferon responsive protein
4	7494391CD1	g13506808	5.70E-42	[Mus musculus] thymic stromal co-transporter Chen, C., et al., Biochim. Biophys. Acta 1493:159-169 (2000)
7	5202657CD1	g3006139	7.90E-42	[Schizosaccharomyces pombe] hypothetical zf-C3HC4 zinc finger protein
15	1661472CD1	g6899934	1.90E-24	[Arabidopsis thaliana] putative zinc-finger protein
17	2159545CD1	g4650844	3.40E-120	[Homo sapiens] Kelch motif containing protein
18	8560269CD1	g18252514	1.00E-123	[Homo sapiens] hepatocellular carcinoma-associated antigen HCA557b
20	6778214CD1	g57671	8.90E-15	[Rattus norvegicus] ribonuclease inhibitor Kawanomoto, M., et al., Biochim. Biophys. Acta 1129:335-338 (1992)
22	2804937CD1	g6650377	0	[Mus musculus] pecanex 1
23	7494915CD1	g2072953	1.60E-34	[Homo sapiens] putative p150 Sassaman, D.M. et al. (1997) Nature Genet. 16:37-43
24	2073751CD1	g18652658	4.00E-17	[Schmidtea mediterranea] myosin heavy chain A
31	2597459CD1	g7243777	2.90E-50	[Drosophila melanogaster] Diablo
35	2804990CD1	g1196425	4.60E-28	[Homo sapiens] envelope protein Cohen, M. et al. (1985) Virology 147:449-458
36	168571CD1	g3093476	6.90E-90	[Homo sapiens] EVI-5 homolog Liao, X. et al. (1997) Oncogene 14:1023-1029
39	2227040CD1	g1799568	1.50E-80	[Homo sapiens] stac Suzuki H., et al. (1996) Biochem. Biophys. Res. Commun. 229:902-909
40	4346130CD1	g15625572	0	[Homo sapiens] centaurin delta1
42	7472392CD1	g12853030	3.00E-81	[Mus musculus] Cyclic nucleotide-binding domain containing protein~data source:Pfam, source key:PF00027, evidence:ISS~putative
43	4028960CD1	g6063688	4.50E-102	[Homo sapiens] AMMECR1 Vitelli, F., et al., Genomics 55:335-340 (1999)
44	8227004CD1	g11527201	6.70E-250	[Homo sapiens] sporulation-induced transcript 4-associated protein SAPLa Twells, R.C.J., et al., Genomics 72:231-242 (2001)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
47	71351918CD1	g1314316	6.30E-73	[Xenopus laevis] WD-40 motifs; up-regulated by thyroid hormone in tadpoles Brown, D.D., et al., Proc. Natl. Acad. Sci. U.S.A. 9:1924-1929 (1996)
52	2321054CD1	g15278367	4.00E-52	[Homo sapiens] Similar to fasciculation and elongation protein zeta 2 (zygin II)
55	7654832CD1	g15420869	0	[Mus musculus] ankyrin repeat-containing SOCS box protein 5 Kile, B.T., et al., Mol. Cell. Biol. 21:6189-6197 (2001)

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	2867236CD1	246	S99 S212 T11 T26 T73 T111 T131	N129 N201		
2	1294096CD1	325	S6 S25 S54 S142 S151 S162 S204 S304 S316 T47 T51 T252 T267 Y119	N254	K+ channel tetramerisation domain: D34-Q134	HMMER_PFAM
3	7238537CD1	376	S34 S52 S96 S106 S113 S118 S264 S359 S373 T38 T231 T286	N152 N319	Cytosolic domain: S321-A376 Transmembrane domain: T298-V320 Non-cytosolic domain: M1-V297	TMHMMER
4	7494391CD1	461	S48 S70 S93 S233 S253 S440 S454 S459 T166 Y39	N38 N46 N53 N143 N229 N251	Sugar transport proteins signature 1: L91-S107	MOTIFS
					Cytosolic domain: M1-V6, D96-K101, D157-R168, G219-F258, S319-D324, K374-L384, K437-R461 Transmembrane domain: E7-Y29, I78-S95, F102-F124, S134-V156, I169-I191, F196-L218, L259-F281, F296-F318, I325-T347, M351-S373, F385-Y407, F417-V436 Non-cytosolic domain: R30-D77, A125-A133, R192-G195, I282-V295, T348-M350, S408-G416	TMHMMER
5	6451054CD1	168	S26 S28 S47 S80 S85 S98 S113 T96	N50		

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
6	7494592CD1	832	S201 S363 S470 S489 S557 S580 S658 S702 S785 S799 S818 T166 T228 T324 T352 T366 T371 T399 T545 T659 T734 Y430	N586		
7	5202657CD1	393	S28 S135 S180 S190 Y35 Y218 Y252	N189	PROTEIN C29A3.03C CHROMOSOME II ZINCFINGER NUCLEAR DNABINDING CODED FOR BY PD024560: E158-F393	BLAST_PRODOM
8	2013529CD1	280	S41 S77 S138 S144 S159 S183 S193 S235 S252 S257 S272 T20 T155 T158 T178	N232		
9	3841351CD1	344	S5 S14 S15 S33 S38 S169 T130 T173 T271 T276	N217		
10	152116CD1	405	S139 S163 S292 S322 S346 S361 T10 T70 T151 T325	N147		
11	2381031CD1	185	S52 S56 S130 S134 S143 S145 S181 T60	N121 N141		

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
12	2511371CD1	463	S8 S261 S300 S330 T47 T88 T166 T313		Trp-Asp (WD) repeat protein BL00678: S210-W220	BLIMPS_BLOCKS
					WD domain, G-beta repeat: C186-R221, P343-D376, P254-K290	HMMER_PFAM
13	8068623CD1	403	S85 S106 S164 S389 T324 T364		TBC domain: E57-C268	HMMER_PFAM
					Cytosolic domain: R261-T364 Transmembrane domain: L238-Y260, N365-V387 Non-cytosolic domain: M1-R237, K388-P403	TMHMMER
14	677977CD1	574	S45 S57 S65 S79 S80 S218 S234 S246 S340 S376 T53 T58 T519	N214	signal_cleavage: M1-A22	SPSCAN
					Cell attachment sequence: R203-D205	MOTIFS
15	1661472CD1	731	S9 S100 S121 S156 S160 S219 S264 S339 S340 S393 S422 S450 S590 S625 S669 S672 T85 T222 T359 T377 T491 T561 T648 Y23	N335	PROTEIN ZINC FINGER CONSERVED CHROMOSOME IV COSMID CODED FOR BY C PD043678: D4-R101	BLAST_PRODOM
					PROLINE-RICH PROTEIN DM03894|P05142|1-134: V460-P550	BLAST_DOMO
					Zinc finger, C2H2 type, domain: C16-H37	MOTIFS
16	1748508CD1	299	S75 S117 S166 S206 T226	N42 N110		

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
17	2159545CD1	620	S7 S32 S34 S127 S218 S297 S307 S354 S475 S522 S531 T70 T106 T311 T316 T390 T401 T448 T542 Y306	N29	BTB/POZ domain: K51-F164	HMMER_PFAM
					Kelch motif: R355-N400, R496-T541, K402-N447, P449-G494, P543-S588, S307-N353	HMMER_PFAM
					PROTEIN REPEAT MATRIX RING CANAL KELCH R12E2.1 C47D12.7 KIAA0132 KIAA0469 PD001473: P166-L294	BLAST_PRODOM
					POZ DOMAIN DM00509|Q04652|131-335: E65-L238 DM00509|A45773|130-334: E65-L238 DM00509|P21073|1-198: F60-Q243	BLAST_DOMO
18	8560269CD1	218	S68 S133 T55 T91 T93 T151	N28 N165	Leucine zipper pattern: L141-L162	MOTIFS
19	8710302CD1	427	S22 S38 T118 T368 Y331		Cytosolic domain: R90-T115 Transmembrane domain: E67-L89, I116-F138 Non-cytosolic domain: M1-P66, R139-S427	TMHMMER
20	6778214CD1	612	S50 S191 S232 S359 S370 S388 S435 S473 S505 S507 S509 S511 S590 S607 T47 T72 T144 T439 T478	N70 N168 N228 N275 N360 N416		

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
21	258383CD1	458	S22 S38 S120 T149 T399 Y362		Cytosolic domain: R90-K95, R170-Q299, S355-S458 Transmembrane domain: E67-L89, F96-L115, I147-F169, L300-R322, I332-A354 Non-cytosolic domain: M1-P66, I116-T146, N323-I331	TMHMMER
22	2804937CD1	1451	S14 S28 S168 S222 S367 S369 S483 S490 S649 S691 S768 S787 S821 S822 S884 S885 S923 S944 S968 S978 S1045 S1093 S1111 S1242 S1247 S1262 S1267 S1330 S1348 S1372 S1392 S1397 S1421 T637 T832 T977 T1438 Y446 Y562 Y635 Y845 Y1355	N201 N265 N689 N830 N1092 N1172 N1182 N1344 N1370	Cytosolic domain:M1-N116, R200-V219, G293-H303, Q399-Y404 Transmembrane domain: V117-F139, I177-S199, F220-L242, L270-Y292, I304-S323, L376-L398, V405-P422 Non-cytosolic domain: R140-I176, L243-S269, R324-D375, Q423-V1451	TMHMMER
					PROTEIN COSMID 30B8 KIAA0435 B0511.12 PECANEX DEVELOPMENTAL NEUROGENESIS TRANSMEMBRANE GLYCOPROTEIN PD014553: N636-G1198	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
22 (cont.)					PROTEIN B0511.12 COSMID 30B8 PECANEX DEVELOPMENTAL NEUROGENESIS TRANSMEMBRANE GLYCOPROTEIN REPEAT PD018553: L66-L287, S354-I471, S280-P359, G1148-N1178	BLAST_PRODOM
					PROTEIN COSMID 30B8 PECANEX DEVELOPMENTAL NEUROGENESIS TRANSMEMBRANE GLYCOPROTEIN REPEAT B0511.12 PD025614: W1379-V1444, P1179-P1207	BLAST_PRODOM
23	7494915CD1	184	S9 T24 T30 T36 T66	N22	Reverse transcriptase (RNA-dependent DNA polymerase): R86-S159	HMMER_PFAM
					Cytosolic domain:M1-L153 Transmembrane domain:F154-V176 Non-cytosolic domain:K177-H184	TMHMMER
					DNA RNADIRECTED POLYMERASE PUTATIVE P150 TRANSCRIPTASE REVERSE PROTEIN L1 SEQUENCE PD002004: Q25-L88	BLAST_PRODOM
					TRANSCRIPTASE; REVERSE; ORF2; ENCODE; DM01377|P08548|132-516: T24-M100 DM01377|P08547|132-516: Q25-M100 DM01377|I38588|130-517: Q25-M100 DM01377|S16783|1-259: Q18-M100	BLAST_DOMO
24	2073751CD1	407	S179 S273 S283 T61 T129 T213 T294 T306 T376 T379 T386 Y22	N247	PROTEIN COILED COIL CHAIN MYOSIN REPEAT HEAVY ATPBINDING FILAMENT HEPTAD PD000002: E128-E366 (Pvalue=2.3e-10)	BLAST_PRODOM
					Leucine zipper pattern: L301-L322	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
25	3178841CD1	261	S40 S56 S70 S140 S144 T234		HYPOTHETICAL 32.0 KD PROTEIN C09F5.2 IN CHROMOSOME III TRANSMEMBRANE PD128096: V45-V183	BLAST_PRODOM
					Cytosolic domains: 1-72, 123-155, 227-261 Transmembrane domains: 79-90, 101-122, 156-178, 204-226 Non-cytosolic domains: 91-99, 179-203	TMHMMER
26	3674807CD1	209	S197 T17 T187	N132	Signal peptide: M41-G90	SPSCAN
27	1794922CD1	333	S6 S17 S39 S65 S140 S174 S212 S241 T31 T217 T310 Y261	N63 N295		MOTIFS
28	1795509CD1	257	T43 T168 Y105		COSMID E04F6 PD132304: F72-C254	BLAST_PRODOM
					Cell attachment sequence: R78-D80	MOTIFS
29	2017180CD1	293	S5 S120 S135 S155 S251 S255 S279 S288 T27 T42 T107 T142	N273 N287		MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
30	219442CD1	598	S28 S66 S71 S102 S149 S163 S180 S186 S191 S215 S228 S263 S283 S287 S361 S365 S406 S439 S543 S561 S571 S584 T45 T67 T172 T185 T276 T290 T337 T472 T492 T545	N195 N220 N227		MOTIFS
31	2597459CD1	470	S3 S8 S23 S78 S177 S196 S201 S359 S377 S451 T157 T279 Y101 Y292 Y309	N25 N155 N325	BTB/POZ domain: D26-L139	HMMER_PFAM
					Kelch motif: G335-K385, C387-D433	HMMER_PFAM
					BTB Domain PF00651: A55-F67	BLIMPS_PFAM
					PROTEIN DNA-BINDING ZINC FINGER METAL BINDING NUCLEAR ZINC FINGER TRANSCRIPTION REGULATION CHROMOSOME PD000632: Q16-L139	BLAST_PRODOM
					POZ DOMAIN DM00509 Q04652|131-335: S21-N218 A45773|130-334: S21-N218 P21073|1-198: S23-E216 S59069|1-171: H24-L135	BLAST_DOMO

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
32	2783863CD1	311	S39 S101 S133 T136	N68	PROTEIN CHROMOSOME READING FRAME ORF TRANSMEMBRANE COSMID D8035.34P XV YOL002C PD005362: N68-S301	BLAST_PRODOM
					Cytosolic domains: 1-73, 129-140, 196-201, 257-275 Transmembrane domains:74-96, 106-128, 141-163, 173-195, 202-224, 239-256, 276-298 Non-cytosolic domains: 97-105, 164-172, 225-238, 299-311	TMHMMR
					MEMBRANE; YOL002C; CHROMOSOME; C30D11.11; DM02642 Q09749|49-323: T32-K302 Q09910|169-441: T32-V283 S62569|169-441: T32-V283 S61982|50-325: Y31-V283	BLAST_DOMO

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
33	2902971CD1	894	S9 S14 S17 S75 S86 S97 S113 S114 S115 S141 S161 S169 S223 S267 S277 S293 S313 S327 S351 S372 S420 S431 S439 S454 S489 S536 S554 S580 S598 S705 S759 S792 S806 S868 T176 T278 T291 T323 T346 T409 T438 T581 T629 T635 T723 Y310	N106 N312 N596 N757	PROTEIN CHROMOSOME C30D11.09 I B0361.1 III PD033465: E673-K786	BLAST_PRODOM
					SPAC30D11.09; DM04663 Q10945|1-144: I662-K786 Q09909|388-532: M672-W796 S62567|388-532: M672-W796	BLAST_DOMO
					Cell attachment sequence: R425-D427	MOTIFS
					ATP/GTP-binding site motif A (P-loop): A622-T629	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
34	368660CD1	653	S37 S69 S79 S135 S225 S290 S303 S475 S482 S483 S512 S515 S516 S520 S582 S645 T373 T384 T414 T448 T453 T476 T601 T629	N288 N508 N542	TPR Domain: V341-H374, A307-N340, C382-F415	HMMER_PFAM
					Cell attachment sequence: R125-D127	MOTIFS
35	2804990CD1	144	T16	N17	Cytosolic domain: 137-144 Transmembrane domain: 114-136 Non-cytosolic domain: 1-113	TMHMMR
36	168571CD1	424	S88 S123 S124 S151 S179 S194 S258 S269 S364 S365 S386 S407 S409 T92	N122 N311	signal_cleavage: M1-L24	SPSCAN
					EVI5 HOMOLOG TRUNCATED EVI5 ECOTROPIC VIRAL INTEGRATION SITE COSMID F01G12 PD075221: E84-P180	BLAST_PRODOM
					Protein kinases ATP-binding region signature: L370-K401	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
37	1286391CD1	1351	S187 S266 S311 S320 S359 S360 S423 S486 S496 S556 S633 S657 S704 S884 S1000 S1204 S1262 S1274 S1296 S1345 T25 T74 T95 T124 T159 T230 T329 T392 T463 T642 T691 T940 T1015 T1079 T1215	N215 N235 N358 N543 N1293	Integrase core domain: R1062-L1130	HMMER_PFAM
					POL POLYPROTEIN DM00159|S08405|760-943: H1009-A1139 DM00140|S52564|1-364: Y517-K670	BLAST_DOMO
					Leucine zipper pattern: L837-L858	MOTIFS
38	2007684CD1	78				
39	2227040CD1	411	S5 S39 S56 S84 S112 S122 S168 S198 S224 S232 S234 S246 S249 S325 S372 S396 T82 T188 T207 T286	N172 N218	Phorbol esters/diacylglycerol binding domain: H111-C161	HMMER_PFAM
					SH3 domain: Y295-V349	HMMER_PFAM
					Phorbol esters / diacylglycerol binding domain proteins BL00479: H111-G133, Q137-C152	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
39 (cont.)					Phorbol esters / diacylglycerol binding domain: R120-S213	PROFILESCAN
					Diacylglycerol/phorbol-ester binding signature PR00008: H158-R170, V108-S122, C124-G133, Q137-V148	BLIMPS_PRINTS
					STAC PD027304: Q347-I411 PD032205: G163-G245	BLAST_PRODOM
					SRC HOMOLOGY 3 (SH3) DOMAIN DM00025|S61138|55-108: Y297-Q347	BLAST_DOMO
					Cytochrome c family heme-binding site signature: C124-Q129	MOTIFS
					Phorbol esters / diacylglycerol binding domain: H111-C161	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
40	4346130CD1	1704	S3 S165 S295 S323 S346 S350 S378 S416 S424 S480 S677 S683 S696 S699 S739 S781 S834 S911 S932 S977 S1124 S1165 S1172 S1194 S1211 S1404 S1428 S1435 S1476 S1477 S1487 S1510 S1582 S1593 S1604 T34 T83 T122 T155 T183 T190 T196 T312 T314 T357 T409 T411 T417 T437 T534 T548 T643 T663 T681 T791 T808 T919 T1011 T1096 T1105 T1215 T1279 T1369 T1423 T1524 T1639 T1656	N42 N167 N205 N231 N327 N694 N697 N943 N1009 N1572	Putative GTP-ase activating protein for Arf: Y685-L807	HMMER_PFAM
					PH domain: S1435-H1537, K483-K574, P588-A679, I879-V1003	HMMER_PFAM
					RhoGAP domain: P1129-E1282	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
40 (cont.)					SAM domain (Sterile alpha motif): V4-L68	HMMER_PFAM
					HIV Rev interacting protein signature PR00405: N697-C716, C716-K733, V484-V505	BLIMPS_PRINTS
					PROTEIN GTPASE DOMAIN AC PD00930: P1129-G1154, L1232-L1272	BLIMPS_PRODOM
					PROTEIN ZINC FINGER NUCLEAR DNA BINDING PUTATIVE GTPASE ACTIVATING FACTOR CHROMOSOME REPEAT PD002425: N694-E775	BLAST_PRODOM
					PROTEIN GTPASE DOMAIN SH2 ACTIVATION ZINC 3 KINASE SH3 PHOSPHATIDYLINOSITOL REGULATORY PD000780: V1128-E1282	BLAST_PRODOM
					PH DOMAIN DM00470 |S54307|1621-1845: K1125-E1301 |P34588|1-285: K1125-N1291 |A49307|566-842: T1096-Q1274 |P15882|109-331: I1107-T1275	BLAST_DOMO
41	55117040CD1	243	S13 S38 S53 S59 S75 T45 T171	N50 N203	Ankyrin repeat: Q135-N167, K168-E200, Y201-Q230	HMMER_PFAM
					Aldehyde ferredoxin oxid PF01314: A110-V122, R73-L103, A82-K114	BLIMPS_PFAM
42	7472392CD1	248	S7 S57 S112 S137 S157 S175 S200 T12 T180			
43	4028960CD1	310	S59 S277 S283 T98 T141 T144 T184 T199 T269	N36 N63 N93 N97 N297		

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
44	8227004CD1	838	S11 S133 S202 S214 S304 S315 S402 S417 S503 S582 S599 S604 S616 S673 S687 S696 S700 S705 S742 S817 T15 T23 T171 T233 T274 T278 T282 T390 T411 T413 T472 T507 T512 T608 T614 T711 T751 T789	N79 N210 N231 N261 N302 N358 N409 N542 N552 N558 N647	PROTEIN SIT4-ASSOCIATING PHOSPHORYLATION CELL CYCLE SAP155 KIAA0685 SAP185 SAP190 SAP4 PD014556: D68-E284	BLAST_PRODOM
					SIT4-ASSOCIATING PROTEIN SAP190 DM03002|P40856|222-821: M92-N363	BLAST_DOMO
					SIT4-ASSOCIATING PROTEIN SAP185 DM03002|P36123|229-825: E98-N363	BLAST_DOMO
45	3044763CD1	408	S25 S62 S68 S103 S184 S286 S330 S344 S377 S392 T98 T131 T158 T213	N257 N342	Leucine zipper pattern: L261-L282	MOTIFS
					Cytosolic domains: M1-G229, R290-R301, S377-S408 Transmembrane domains: L230-G252, T267-D289, N302-L324, I354-L376 Non-cytosolic domains: D253-K266, S325-Q353	TMHMMER
46	4044519CD1	101	T89		Signal_cleavage: M1-S15	SPSCAN
					Signal Peptide: M1-S19, M1-T16, M31-S59	HMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
46 (cont.)					Cytosolic domains: M1-M31, E82-L101 Transmembrane domains: V32-L54, L64-L81 Non-cytosolic domain: N55-S63	TMHMMER
47	71351918CD1	256	S26 S107 S174 S194 S202 T35 T63 T222		WD domain, G-beta repeat: C32-D68, Q136-D171, V179-Q212	HMMER_PFAM
					Non-cytosolic domain: M1-L256	TMHMMER
					Trp-Asp (WD-40) repeats signature: L44-A92	PROFILESCAN
					Trp-Asp containing, G-protein WD-40 repeats GENE 16 PD106308: R71-L256	BLAST_PRODOM
					Trp-Asp (WD) repeats signature: V55-C69	MOTIFS
48	8109363CD1	104			Non-cytosolic domain: M1-G104	TMHMMER
49	1272746CD1	855	S27 S93 S159 S184 S245 S287 S329 S593 S708 S726 S793 S811 S827 T18 T26 T35 T45 T128 T200 T202 T319 T352 T438 T470 T601 T642 T807	N132 N157 N254 N312 N325 N486 N592 N615 N623 N659 N822	Non-cytosolic domain: M1-S855	TMHMMER
50	1839974CD1	427	S254 S277 S321 S334 S361 S365 T50 T182 T247 Y329	N73 N258 N405	Fibronectin type III domain: L276-G364	HMMER_PFAM
					Non-cytosolic domain: M1-I427	TMHMMER
					Fibronectin type III repeat signature PR00014: S393-P402, G406-W416	BLIMPS_PRINTS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
51	1877336CD1	800	S58 S74 S252 S277 S283 S306 S351 S380 S415 S424 S437 S439 S446 S491 S521 S534 S539 S605 S786 T208 T302 T313 T488 T553 T572 T747 T781 Y331	N272 N281 N411 N785	Non-cytosolic domain: M1-F800	TMHMMER
					FIBRILLAR COLLAGEN CARBOXYL-TERMINAL DM00019|S42886|221-377: G112-N237	BLAST_DOMO
52	2321054CD1	107	S87 T17 T62		Signal_cleavage: M1-A43	SPSCAN
					Non-cytosolic domain: M1-E24 Transmembrane domain: A25-A43 Cytosolic domain: K44-E107	TMHMMER
53	2796034CD1	522	S10 S49 S77 S191 S195 S204 S212 S335 S339 S344 S361 S371 S397 S443 S453 T72 T125 T138 T146 T228 T241 T245 T267 T318 T363 T496	N59 N85 N120 N487 N494	Non-cytosolic domain: M1-Q522	TMHMMER
54	4413112CD1	305	S28 S57 S98 S122 S238 S294	N44 N55 N130 N148	Signal_cleavage: M1-S20	SPSCAN
					Signal Peptide: M1-S20	HMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
54 (cont.)					Non-cytosolic domain: M1-T165 Transmembrane domain: I166-W188 Cytosolic domain: R189-H305	TMHMMER
					Leucine-rich repeat signature PR00019: L84-L97, V63-L76	BLIMPS_PRINTS
55	7654832CD1	329	S2 S67 S150 S244 S277 T121		Signal_cleavage: M1-G46	SPSCAN
					Cytosolic domain: M1-T20 Transmembrane domain: I21-I43 Non-cytosolic domain: V44-R329	TMHMMER
					Ank repeat: Y232-T264, D102-I134, A69-L101, D135-C167, S170-P199, H200-K231	HMMER_PFAM
56	7503849CD1	236	S34 S52 S96 S106 S113 S118 T38 T231	N152	Signal peptide: M18-A80	SPSCAN

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
57/2867236CB1/ 1485	1-222, 1-461, 1-464, 1-572, 4-516, 7-279, 12-266, 16-285, 23-278, 24-281, 24-522, 30-320, 31-309, 31-316, 34-172, 34-327, 34-639, 42-282, 46-218, 46-335, 51-290, 51-342, 141-404, 141-405, 143-579, 182-471, 208-654, 208-684, 211-695, 217-453, 321-836, 335-910, 381-933, 390-996, 475-747, 475-762, 482-1009, 537-1153, 591-970, 717-993, 796-1090, 797-1382, 808-1485, 963-1280
58/1294096CB1/ 6176	1-210, 1-383, 1-416, 1-508, 1-516, 1-6176, 33-443, 33-570, 228-695, 228-1034, 365-486, 495-1170, 511-962, 574-961, 647-975, 794-1062, 959-1554, 1000-1160, 1009-1246, 1009-1497, 1034-1327, 1075-1207, 1111-1292, 1170-1258, 1250-1842, 1260-1540, 1261-1835, 1288-1927, 1292-1877, 1352-1518, 1417-1679, 1438-2085, 1446-1787, 1542-1841, 1697-1891, 1860-2146, 1872-1935, 1971-2491, 2021-2670, 2028-2623, 2108-2393, 2173-2698, 2174-2729, 2179-2427, 2192-2528, 2194-2469, 2196-2775, 2198-2432, 2205-2471, 2206-2456, 2208-2458, 2211-2451, 2212-2465, 2226-2443, 2230-2787, 2242-2498, 2270-2481, 2281-2922, 2296-2933, 2310-2564, 2320-2575, 2327-2614, 2329-2701, 2337-2663, 2350-2472, 2384-2671, 2384-3130, 2397-2558, 2397-2976, 2418-2677, 2450-2878, 2455-2608, 2470-2665, 2479-2725, 2480-2676, 2484-2768, 2496-3042, 2506-2905, 2509-2739, 2509-3040, 2512-2772, 2563-3211, 2603-2902, 2665-2928, 2667-3092, 2673-3232, 2706-3159, 2711-3168, 2719-2972, 2726-3166, 2729-2927, 2738-3165, 2743-2840, 2744-3166, 2750-2997, 2763-3037, 2779-3309, 2780-3064, 2797-3413, 2822-3091, 2829-3021, 2836-3285, 2843-3169, 2848-3438, 2852-3032, 2852-3209, 2856-3109, 2861-3546, 2881-3106,
	2881-3375, 2890-3411, 2907-3364, 2910-3158, 2964-3435, 2993-3483, 2993-3597, 3009-3277, 3051-3494, 3056-3286, 3069-3279, 3069-3583, 3071-3355, 3071-3660, 3098-3530, 3103-3348, 3126-3582, 3134-3337, 3140-3709, 3145-3396, 3174-3319, 3199-3388, 3199-3854, 3201-3770, 3204-3779, 3222-3627, 3250-3667, 3284-3920, 3308-3625, 3336-3570, 3336-3594, 3336-3598, 3336-3601, 3336-3615, 3336-3617, 3336-3629, 3336-3631, 3336-3636, 3336-3669, 3342-3867, 3342-3957, 3342-3961, 3361-3965, 3370-3645, 3381-3575, 3381-3580, 3420-3832, 3435-4105, 3460-3723, 3513-4121, 3519-3784, 3521-3985, 3542-3680, 3544-3775, 3559-4065, 3569-4170, 3576-3824, 3601-3810, 3625-3753, 3633-3852, 3639-4169, 3644-3940, 3651-4246, 3673-4260, 3682-3853, 3716-4031, 3720-4094, 3720-4186, 3720-4194, 3721-4281, 3728-4206, 3728-4340, 3728-4427, 3749-4132, 3750-4001, 3750-4297, 3751-3880, 3759-4040, 3768-4266, 3778-4243, 3781-4065, 3783-4453, 3789-4049, 3796-4095, 3803-4392, 3820-4049, 3820-4052, 3820-4106, 3835-4309, 3841-4163, 3841-4185, 3852-4098, 3853-4437, 3865-4082, 3865-4389, 3868-4165, 3892-4082, 3898-4136, 3899-4211, 3937-4467, 3940-4545, 3945-4098, 3954-4496, 3963-4183,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
58 (cont.)	3973-4467, 3978-4383, 3987-4288, 3987-4535, 3991-4468, 3993-4595, 4004-4256, 4009-4482, 4011-4475, 4013-4496, 4014-4292, 4017-4496, 4018-4481, 4020-4496, 4027-4496, 4028-4496, 4033-4480, 4037-4496, 4049-4496, 4050-4479, 4050-4496, 4054-4481, 4055-4481, 4057-4496, 4063-4481, 4067-4496, 4071-4480, 4077-4496, 4079-4475, 4080-4480, 4088-4355, 4096-4480, 4099-4544, 4099-4580, 4107-4481, 4125-4421, 4130-4448, 4131-4486, 4136-4236, 4144-4481, 4145-4780, 4152-4421, 4162-4479, 4169-4480, 4169-4486, 4178-4464, 4180-4484, 4185-4451, 4185-4544, 4185-4769, 4187-4402, 4192-4361, 4198-4476, 4224-4405, 4224-4445, 4237-4483, 4241-4496, 4249-4481, 4278-4715, 4283-4397, 4327-4587, 4370-4630, 4373-4627, 4402-4496, 4409-4717, 4449-4731, 4451-4750, 4467-4618, 4468-4820, 4478-4720, 4478-5043, 4490-4908, 4492-4738, 4493-4740, 4497-5079, 4499-4646, 4503-4776, 4506-4737, 4525-4787, 4527-4755, 4552-5003, 4556-4852, 4561-4822, 4562-5252, 4564-5277, 4572-4875, 4574-4841, 4591-5200, 4603-4856, 4604-4867, 4605-4849, 4605-4988, 4606-4864, 4610-4892, 4614-4818, 4614-5195, 4625-4875, 4625-4881, 4644-5326, 4649-4940, 4655-4824, 4657-5067, 4672-4965, 4677-5283,
	4690-4934, 4694-4916, 4698-4976, 4700-4949, 4707-5179, 4715-4975, 4717-5277, 4727-4987, 4732-5006, 4732-5009, 4743-4962, 4744-5327, 4749-5009, 4770-5078, 4780-5025, 4789-4995, 4798-5040, 4804-5290, 4823-4912, 4823-5407, 4824-4943, 4825-5077, 4829-5103, 4833-5086, 4855-5387, 4874-5558, 4912-5366, 4916-5718, 4919-5475, 4987-5562, 4991-5497, 5003-5554, 5009-5551, 5015-5373, 5059-5359, 5087-5585, 5100-5412, 5117-5368, 5118-5354, 5123-5387, 5132-5378, 5132-5443, 5135-5600, 5136-5360, 5140-5388, 5159-5473, 5165-5390, 5166-5424, 5169-5429, 5184-5409, 5189-5356, 5190-5680, 5205-5467, 5205-5505, 5206-5507, 5206-5759, 5213-5491, 5223-5507, 5227-5503, 5233-5953, 5236-5506, 5236-5734, 5238-5493, 5238-5507, 5248-5499, 5249-5575, 5255-5376, 5267-5525, 5278-5587, 5279-5566, 5282-5555, 5286-5424, 5289-5901, 5290-5505, 5298-5566, 5301-5559, 5353-5640, 5354-5549, 5354-5597, 5354-5601, 5389-5912, 5412-5957, 5479-6039, 5544-5877, 5547-6168, 5574-6129, 5591-5846, 5591-5990, 5591-6127, 5594-6062, 5599-6176, 5601-5849, 5603-6069, 5609-5864, 5609-5868, 5609-5925, 5609-5979, 5609-6133, 5611-5894, 5617-5868, 5619-5882, 5623-5862, 5632-5890, 5638-6176,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
58 (cont.)	5645-5866, 5655-6169, 5659-5853, 5659-6082, 5664-5866, 5664-5911, 5671-5929, 5671-6063, 5671-6068, 5671-6169, 5672-5929, 5673-5914, 5674-6169, 5682-6154, 5688-6169, 5689-5913, 5689-5961, 5689-5973, 5691-6095, 5692-6145, 5696-5948, 5697-6149, 5697-6169, 5699-5987, 5699-5993, 5703-6169, 5704-5982, 5704-6151, 5704-6169, 5709-6166, 5710-5872, 5710-6170, 5712-5984, 5712-5995, 5712-6169, 5713-5872, 5716-6153, 5720-6176, 5729-6169, 5735-6145, 5745-6170, 5746-6169, 5749-5942, 5752-5994, 5752-6169, 5753-6022, 5753-6171, 5754-6170, 5756-5999, 5756-6169, 5756-6174, 5757-6169, 5759-6169, 5759-6176, 5761-6169, 5762-6149, 5763-5996, 5764-6169, 5771-6035, 5771-6174, 5772-5994, 5775-5959, 5776-6016, 5780-6169, 5783-6061, 5783-6169, 5785-6095, 5786-5991, 5786-6113, 5788-6169, 5789-6073, 5803-6071, 5805-6076, 5811-6173, 5812-6169, 5812-6176, 5814-6075, 5814-6142, 5814-6176, 5815-6174, 5816-6169, 5816-6171, 5817-6153, 5818-6169, 5819-6145, 5820-6160, 5820-6173, 5820-6176, 5822-6176, 5824-6031, 5824-6169, 5824-6171, 5827-6149, 5830-6168, 5830-6169, 5832-6017, 5844-6170, 5846-6083, 5846-6095, 5846-6107, 5846-6170, 5850-6169, 5852-6152, 5852-6170,
	5854-6095, 5861-6173, 5868-6169, 5874-6133, 5874-6169, 5881-6122, 5883-6136, 5887-6173, 5928-6162, 5928-6176, 5931-6171, 5932-6133, 5935-6151, 5935-6162, 5935-6163, 5935-6174, 5937-6169, 5941-6166, 5941-6176, 5945-6166, 5957-6150, 5958-6119, 5958-6145, 5958-6154, 5962-6163, 5962-6166, 5963-6139, 5971-6147, 5971-6176, 5973-6145, 5979-6175, 5989-6164, 5996-6137, 6005-6171, 6006-6151, 6006-6176, 6020-6169, 6022-6145, 6027-6142, 6051-6169, 6063-6169, 6087-6166, 6101-6169
59/7238537CB1/ 1944	1-496, 13-568, 15-704, 16-417, 16-532, 17-371, 18-272, 18-325, 19-309, 20-571, 21-215, 25-288, 25-322, 27-288, 32-623, 34-671, 35-287, 40-682, 41-638, 44-642, 54-372, 138-407, 147-380, 161-429, 204-838, 233-497, 233-767, 268-897, 297-829, 305-729, 310-664, 336-894, 338-586, 346-615, 352-938, 354-702, 369-771, 380-1045, 388-837, 394-1041, 428-710, 428-1061, 428-1206, 462-892, 472-896, 494-1232, 496-1077, 511-1063, 519-777, 527-797, 530-772, 531-770, 548-1220, 558-1229, 559-1143, 561-1116, 561-1220, 578-1194, 584-1232, 586-864, 595-861, 603-870, 634-1312, 662-1147, 681-1042, 681-1388, 686-1169, 703-1209, 710-1189, 723-1149, 731-1284, 741-1209, 744-1175, 760-1324, 766-1079, 774-1188, 792-1051, 802-1160, 807-1052, 813-1120, 813-1386, 834-1191, 840-1129, 861-1318, 861-1439, 861-1470, 866-1128, 868-1403, 881-1117, 883-1122, 889-1516, 891-1156, 891-1168, 893-1131, 894-1158, 896-1601, 904-1387, 908-1634, 915-1398, 924-1163, 924-1500, 924-1544, 930-1600, 931-1187, 931-1489, 934-1658, 938-1535, 942-1456, 947-1210, 951-1250, 957-1214, 959-1422, 967-1585, 970-1422, 976-1226, 981-1419, 985-1418, 987-1704, 992-1520, 995-1259, 1003-1269, 1003-1567, 1005-1703, 1006-1585,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
59 (cont.)	1010-1204, 1010-1263, 1010-1623, 1013-1303, 1014-1632, 1019-1354, 1024-1483, 1025-1545, 1031-1662, 1039-1302, 1039-1418, 1039-1593, 1051-1698, 1052-1416, 1059-1300, 1063-1600, 1065-1422, 1066-1404, 1067-1464, 1073-1409, 1073-1593, 1074-1369, 1074-1419, 1075-1544, 1079-1323, 1081-1668, 1086-1544, 1088-1697, 1093-1378, 1099-1544, 1101-1424, 1103-1544, 1110-1460, 1117-1462, 1128-1630, 1133-1703, 1135-1702, 1138-1617, 1140-1703, 1143-1550, 1145-1369, 1147-1544, 1153-1690, 1155-1702, 1157-1698, 1159-1696, 1171-1335, 1178-1752, 1192-1544, 1200-1547, 1209-1540, 1213-1506, 1224-1928, 1236-1704, 1243-1704, 1247-1705, 1256-1703, 1256-1705, 1261-1705, 1263-1561, 1270-1704, 1273-1704, 1288-1705, 1290-1705, 1293-1696, 1293-1944, 1294-1704, 1295-1704, 1297-1704, 1308-1703, 1308-1704, 1311-1568, 1315-1666, 1317-1704, 1318-1705, 1322-1705, 1324-1705, 1327-1627, 1327-1702, 1328-1704, 1332-1704, 1333-1704, 1334-1693, 1337-1696, 1344-1613, 1353-1722, 1355-1650, 1362-1705, 1366-1654, 1366-1660, 1366-1704, 1372-1804, 1380-1704, 1386-1705, 1388-1645, 1389-1704, 1396-1704, 1403-1704, 1426-1682, 1480-1693, 1483-1704, 1494-1717, 1494-1737, 1542-1705
60/7494391CB1/ 1992	1-669, 172-838, 295-743, 295-786, 295-843, 295-931, 295-934, 295-954, 295-1002, 295-1021, 299-774, 299-919, 300-582, 300-1127, 346-1138, 347-974, 350-1163, 414-1050, 430-1203, 435-1167, 439-642, 444-517, 481-1162, 532-1242, 557-1282, 587-1396, 652-1310, 674-1162, 755-1546, 756-1560, 804-1044, 805-1586, 822-1464, 822-1614, 838-1064, 883-1583, 922-1557, 924-1710, 961-1686, 964-1657, 965-1207, 990-1471, 1025-1312, 1025-1665, 1032-1842, 1033-1648, 1072-1643, 1075-1507, 1087-1921, 1099-1602, 1114-1555, 1121-1356, 1121-1475, 1134-1474, 1156-1376, 1156-1825, 1161-1707, 1168-1893, 1190-1814, 1229-1870, 1255-1903, 1258-1769, 1266-1869, 1282-1575, 1300-1806, 1325-1992, 1347-1540

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
61/6451054CB1/ 3906	1-24, 1-27, 1-36, 1-41, 1-42, 1-43, 1-53, 1-55, 1-64, 1-65, 1-71, 1-72, 1-75, 1-80, 1-83, 1-86, 1-88, 1-89, 1-90, 1-92, 1-94, 1-95, 1-109, 1-117, 1-123, 1-130, 1-131, 1-139, 1-140, 1-170, 1-172, 1-186, 1-189, 1-190, 1-191, 1-193, 1-399, 5-189, 6-186, 7-189, 10-159, 14-151, 22-193, 29-193, 38-192, 38-193, 41-112, 44-189, 59-90, 71-193, 78-193, 79-193, 86-193, 87-193, 91-192, 96-192, 122-193, 126-193, 127-193, 133-186, 139-192, 151-192, 153-191, 158-193, 162-193, 192-680, 192-728, 433-998, 433-1007, 433-1039, 433-1040, 444-668, 453-1350, 503-782, 573-1070, 573-1166, 573-3748, 633-984, 633-1139, 639-1306, 666-1172, 683-950, 684-1132, 703-1346, 755-1518, 790-1294, 814-1423, 867-1641, 869-1468, 909-1501, 996-1644, 1009-1598, 1025-1654, 1090-1117, 1103-1320, 1103-1610, 1105-1535, 1105-1550, 1131-1438, 1136-1465, 1153-1704, 1158-1465, 1177-1419, 1178-1253, 1198-1709, 1201-1552, 1210-1552, 1224-1839, 1227-1486, 1239-1540, 1240-1540, 1249-1788, 1269-1821, 1292-1804, 1317-1499, 1332-1556, 1426-2015, 1465-1953, 1471-1707, 1487-1693, 1509-2141, 1519-1798, 1520-1609, 1570-1912, 1593-2047, 1621-1891, 1631-2218, 1723-1882, 1726-2213, 1762-2341, 1764-2348, 1764-2431, 1765-2431, 1785-2431,
	1816-2078, 1822-2074, 1843-2044, 1861-2292, 1874-2214, 1891-2776, 1921-2298, 1943-2431, 1969-2494, 1990-2610, 2003-2306, 2011-2255, 2029-2290, 2029-2442, 2078-2299, 2079-2599, 2094-2558, 2121-2320, 2121-2323, 2127-2428, 2128-2336, 2128-2623, 2129-2857, 2136-2731, 2140-2430, 2155-2620, 2191-2846, 2198-2731, 2200-2731, 2206-2460, 2210-2444, 2211-2367, 2217-3127, 2227-2447, 2243-2806, 2320-2898, 2325-2981, 2340-2767, 2354-2493, 2356-2466, 2358-2592, 2358-2740, 2373-2485, 2374-2561, 2374-2857, 2374-2892, 2378-2661, 2380-2652, 2383-2656, 2394-2976, 2396-3005, 2430-3135, 2447-2699, 2447-2706, 2470-3044, 2473-2759, 2474-3022, 2482-3005, 2492-3007, 2498-2986, 2502-2946, 2503-2634, 2510-2706, 2510-3005, 2515-2964, 2516-2747, 2522-3007, 2524-2768, 2527-2656, 2535-3028, 2555-2858, 2555-3132, 2555-3133, 2556-2825, 2559-2819, 2573-2795, 2593-2863, 2595-2857, 2598-2836, 2603-2876, 2603-2926, 2617-3043, 2619-2887, 2620-3067, 2623-2761, 2635-3021, 2646-2908, 2649-3215, 2651-3260, 2652-3135, 2653-3065, 2659-3420, 2665-2946, 2679-3186, 2683-3348, 2684-3235, 2686-3166, 2687-3162, 2689-3174, 2711-2945, 2716-2983, 2732-2988, 2737-3247, 2739-3202,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
61 (cont.)	2776-3348, 2783-3290, 2789-3052, 2790-3247, 2796-3045, 2799-3054, 2802-3423, 2830-3094, 2834-3460, 2846-3098, 2847-3046, 2862-3145, 2866-3107, 2867-3029, 2869-3162, 2871-3253, 2871-3410, 2875-3131, 2891-3193, 2893-3193, 2900-3149, 2900-3494, 2901-3485, 2914-3137, 2922-3185, 2922-3381, 2925-3444, 2927-3495, 2932-3584, 2947-3420, 2957-3233, 2959-3210, 2966-2998, 2991-3259, 2994-3597, 2997-3580, 3004-3264, 3013-3571, 3016-3181, 3023-3228, 3025-3299, 3025-3348, 3025-3662, 3026-3224, 3046-3308, 3051-3315, 3053-3526, 3057-3562, 3059-3739, 3063-3300, 3079-3640, 3081-3742, 3096-3564, 3104-3748, 3105-3738, 3116-3737, 3119-3680, 3121-3358, 3127-3307, 3127-3647, 3128-3432, 3128-3701, 3129-3295, 3133-3513, 3145-3733, 3146-3739, 3156-3365, 3156-3417, 3156-3427, 3156-3616, 3156-3738, 3175-3733, 3184-3635, 3188-3418, 3190-3742, 3193-3386, 3200-3407, 3200-3697, 3214-3457, 3214-3742, 3216-3738, 3227-3676, 3235-3691, 3246-3742, 3246-3743, 3247-3746, 3255-3531, 3255-3748, 3256-3742, 3263-3491, 3263-3751, 3269-3536, 3274-3345, 3277-3396, 3277-3549, 3291-3748, 3292-3742, 3294-3748, 3296-3327, 3296-3703, 3296-3760, 3297-3742, 3298-3748, 3300-3548,
	3300-3559, 3300-3743, 3300-3748, 3303-3748, 3306-3746, 3311-3748, 3314-3748, 3315-3748, 3321-3756, 3322-3755, 3322-3760, 3331-3754, 3333-3760, 3342-3749, 3343-3746, 3344-3747, 3345-3616, 3346-3742, 3351-3750, 3360-3746, 3363-3748, 3371-3713, 3371-3748, 3372-3524, 3375-3741, 3377-3746, 3378-3748, 3382-3740, 3389-3748, 3395-3674, 3397-3742, 3409-3664, 3410-3746, 3413-3746, 3417-3688, 3419-3711, 3423-3746, 3428-3760, 3440-3746, 3443-3748, 3450-3746, 3450-3748, 3455-3746, 3461-3732, 3469-3760, 3480-3681, 3482-3543, 3496-3735, 3496-3748, 3499-3732, 3504-3733, 3504-3735, 3504-3747, 3504-3760, 3519-3742, 3528-3748, 3535-3748, 3585-3748, 3603-3735, 3609-3748, 3612-3748, 3614-3906, 3636-3748, 3665-3739
62/7494592CB1/ 3236	1-234, 8-614, 30-557, 30-619, 31-226, 102-429, 127-689, 134-799, 150-400, 186-834, 238-571, 520-1276, 624-1216, 688-1195, 721-851, 792-1518, 821-1162, 835-1154, 905-1626, 916-1626, 922-1626, 943-1222, 953-1499, 1086-1616, 1112-1702, 1246-1665, 1251-1676, 1313-1858, 1441-1854, 1465-1919, 1477-1620, 1642-2180, 1697-1860, 1716-2387, 1778-2128, 1796-2083, 1822-2371, 1846-2158, 1905-2422, 1973-2308, 2075-2187, 2092-2626, 2092-2631, 2094-2362, 2101-2662, 2191-2810, 2243-2655, 2506-2824, 2570-3236, 2572-2827

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
63/5202657CB1/ 1906	1-237, 1-525, 2-270, 3-277, 3-300, 3-557, 3-567, 7-270, 7-275, 8-292, 8-529, 12-207, 12-248, 12-606, 26-274, 26-541, 27-276, 28-572, 29-306, 33-244, 33-330, 76-714, 119-450, 119-451, 125-371, 272-537, 296-556, 324-467, 553-990, 555-813, 555-961, 557-812, 557-814, 571-812, 605-799, 608-860, 656-884, 661-927, 669-931, 712-959, 712-976, 712-992, 713-980, 716-1007, 777-1048, 798-1050, 884-1405, 890-1120, 902-1175, 929-1294, 998-1533, 1018-1432, 1037-1198, 1055-1230, 1061-1563, 1062-1283, 1064-1326, 1066-1342, 1068-1333, 1075-1396, 1079-1305, 1085-1345, 1085-1370, 1094-1347, 1102-1313, 1102-1321, 1102-1658, 1102-1789, 1139-1396, 1151-1648, 1177-1403, 1183-1867, 1198-1894, 1205-1491, 1210-1827, 1226-1563, 1230-1834, 1237-1663, 1259-1874, 1264-1838, 1265-1544, 1268-1882, 1275-1566, 1284-1586, 1296-1516, 1303-1496, 1344-1583, 1346-1906, 1353-1874, 1359-1594, 1366-1748, 1373-1635, 1375-1656, 1375-1851, 1378-1618, 1378-1619, 1393-1594, 1393-1631, 1393-1656, 1398-1751, 1425-1894, 1441-1898, 1446-1872, 1450-1906, 1455-1855, 1459-1891, 1462-1718, 1462-1878, 1467-1864, 1475-1898, 1520-1793, 1520-1890, 1529-1878, 1536-1760, 1536-1823, 1558-1891, 1560-1890, 1587-1792,
	1602-1870, 1606-1851, 1622-1872, 1636-1906, 1697-1906, 1708-1886, 1755-1891, 1776-1874
64/2013529CB1/ 1347	1-153, 1-293, 1-303, 1-322, 1-383, 1-423, 1-469, 1-478, 1-483, 1-484, 1-490, 1-507, 1-560, 1-603, 1-611, 1-627, 1-654, 1-659, 1-660, 2-554, 2-630, 15-313, 15-636, 25-303, 25-328, 48-793, 64-554, 108-793, 183-441, 254-777, 279-567, 282-556, 331-616, 341-1071, 350-604, 350-889, 374-591, 374-665, 413-610, 415-683, 455-1085, 517-848, 517-924, 518-1149, 529-1098, 573-1203, 618-998, 634-1303, 647-1224, 675-1318, 681-895, 681-1040, 681-1057, 681-1077, 681-1090, 681-1091, 681-1094, 681-1199, 681-1201, 681-1205, 681-1215, 681-1223, 681-1303, 681-1312, 684-1127, 684-1308, 751-1347, 761-1347, 771-1347, 816-1312, 908-1312, 916-1334, 929-1310, 1062-1347, 1089-1310, 1128-1312, 1172-1310
65/3841351CB1/ 1854	1-585, 1-1854, 21-594, 29-721, 31-630, 34-302, 46-582, 49-289, 74-312, 474-875, 632-884, 632-1014, 632-1230, 632-1287, 675-1276, 761-1254, 814-1392, 850-1273, 864-1103, 904-1208, 906-1431, 945-1276, 949-1268, 1207-1854, 1208-1608, 1316-1854, 1378-1603, 1410-1800, 1473-1560
66/152116CB1/ 1327	1-568, 74-694, 81-697, 93-579, 105-775, 108-291, 239-902, 468-718, 468-869, 515-761, 579-838, 583-1327, 670-1289, 685-1291, 700-896, 872-1287

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
67/2381031CB1/ 627	1-225, 1-515, 1-604, 2-254, 3-272, 7-185, 7-266, 7-286, 7-290, 7-291, 7-294, 7-296, 8-207, 9-301, 11-307, 11-319, 15-264, 16-250, 17-278, 17-290, 17-295, 17-572, 18-246, 20-373, 22-301, 29-271, 29-287, 32-202, 32-257, 32-273, 32-291, 32-338, 32-343, 35-293, 35-295, 35-348, 37-194, 37-397, 37-452, 43-360, 51-327, 154-435, 155-617, 156-433, 157-627, 158-607, 158-612, 162-378, 162-379, 164-615, 168-479, 169-627, 185-627, 188-606, 189-627, 204-609, 223-607, 233-409, 254-615, 284-556, 292-573, 300-606, 344-529, 366-612, 418-627
68/2511371CB1/ 2564	1-577, 237-1628, 328-691, 328-892, 334-925, 373-975, 496-1062, 557-1106, 686-855, 921-978, 1002-1634, 1053-1594, 1091-1766, 1102-1630, 1209-1827, 1210-1767, 1213-1642, 1213-1926, 1286-1943, 1305-1700, 1316-1945, 1336-1800, 1341-1981, 1341-2008, 1344-1908, 1345-1935, 1349-1899, 1464-2013, 1492-2061, 1503-1967, 1569-2167, 1581-2170, 1662-2343, 1695-2351, 1746-2337, 1746-2400, 1750-2394, 1756-2418, 1772-2316, 1776-2497, 1902-2429, 1924-2453, 1930-2564
69/8068623CB1/ 4134	1-1457, 1292-1820, 1292-1831, 1292-1843, 1292-1908, 1292-1933, 1305-2006, 1306-1967, 1307-1900, 1307-1942, 1313-1790, 1319-1826, 1320-2234, 1322-1920, 1327-1889, 1331-1976, 1332-1610, 1352-1908, 1399-1856, 1439-1929, 1485-1694, 1485-2040, 1501-1939, 1510-2089, 1516-2152, 1516-2156, 1516-2175, 1616-1985, 1695-1933, 1695-1942, 1738-1934, 1810-2017, 1810-2271, 1862-2339, 1882-2248, 1941-2478, 1961-2212, 1961-2214, 1970-2218, 1975-2115, 1989-2719, 2009-2287, 2019-2457, 2019-2568, 2029-2269, 2046-2627, 2051-2697, 2055-2587, 2167-2434, 2185-2435, 2213-2829, 2258-2521, 2336-2525, 2408-2941, 2409-2620, 2430-2714, 2478-3024, 2547-3147, 2670-2933, 2674-2921, 2674-3216, 2788-3037, 2867-3224, 2867-3445, 2908-3173, 2959-3159, 2959-3186, 2998-3252, 2998-3544, 3104-3358, 3195-3383, 3195-3432, 3285-3557, 3310-3502, 3368-3647, 3387-3642, 3410-3598, 3491-3643, 3491-4021, 3499-3753, 3554-4134, 3610-3749, 3629-3859, 3652-3900, 3776-4011
70/677977CB1/ 2329	1-1548, 464-1093, 467-1093, 482-1093, 496-1093, 502-1093, 504-1093, 506-1093, 513-1093, 515-1093, 528-1093, 529-1093, 569-1093, 588-1093, 612-1088, 626-1093, 627-1093, 655-976, 678-1093, 694-1093, 903-1166, 927-1448, 1119-1860, 1161-1830, 1211-1741, 1253-1481, 1374-2024, 1378-1640, 1414-1546, 1414-1548, 1439-1704, 1461-2042, 1467-1985, 1482-1548, 1512-1548, 1516-1548, 1519-1548, 1522-1548, 1548-1636, 1548-1733, 1548-1740, 1548-1763, 1548-1764, 1548-1788, 1548-2001, 1548-2080, 1548-2147, 1548-2195, 1548-2210, 1548-2220, 1548-2329, 1578-1656, 1582-1851, 1585-2146, 1588-1849, 1588-2097

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
71/1661472CB1/ 2960	1-389, 2-389, 4-26, 18-2925, 96-815, 118-777, 333-923, 366-622, 369-907, 482-1135, 536-1004, 573-812, 661-1021, 739-973, 739-1245, 743-1323, 774-1466, 793-1368, 882-1005, 927-1440, 964-1514, 1008-1582, 1035-1321, 1291-1719, 1305-1567, 1305-1721, 1369-1547, 1419-1666, 1419-1795, 1630-1685, 1675-2070, 1766-2529, 1781-2148, 1781-2152, 1781-2164, 1781-2210, 1781-2221, 1781-2275, 1781-2292, 1781-2457, 1782-2446, 1799-2194, 1812-2291, 1839-2037, 1839-2318, 1843-2466, 1883-2459, 1893-2298, 1904-2271, 1935-2144, 1935-2150, 1939-2197, 1941-2200, 1941-2541, 1941-2621, 1942-2327, 1949-2152, 1949-2206, 1994-2214, 1996-2151, 1997-2253, 2004-2298, 2005-2756, 2010-2248, 2028-2282, 2032-2262, 2032-2283, 2038-2258, 2038-2372, 2038-2489, 2039-2723, 2046-2331, 2053-2544, 2055-2802, 2102-2744, 2111-2313, 2127-2345, 2127-2622, 2132-2914, 2140-2373, 2160-2754, 2161-2417, 2164-2392, 2166-2412, 2176-2445, 2203-2384, 2233-2883, 2239-2536, 2239-2548, 2259-2497, 2270-2541, 2272-2520, 2285-2451, 2285-2761, 2306-2855, 2328-2760, 2341-2592, 2349-2619, 2349-2796, 2357-2810, 2372-2948, 2377-2628, 2385-2811, 2391-2811, 2403-2893, 2439-2857, 2443-2920, 2453-2926,
	2468-2925, 2478-2921, 2483-2922, 2488-2960, 2497-2938, 2499-2923, 2500-2917, 2502-2920, 2504-2948, 2511-2923, 2515-2923, 2520-2928, 2543-2926, 2549-2700, 2557-2920, 2568-2923, 2584-2934, 2595-2924, 2597-2914, 2726-2920, 2770-2924, 2773-2924
72/1748508CB1/ 2623	1-523, 57-749, 89-761, 392-889, 473-739, 545-1327, 592-1314, 640-1337, 678-1404, 706-1471, 741-868, 860-1366, 896-1143, 912-1385, 916-1156, 920-1172, 981-1578, 1067-1925, 1097-1796, 1150-1427, 1182-1797, 1240-1957, 1299-1529, 1322-1592, 1322-1603, 1322-1617, 1322-1737, 1322-1931, 1322-1962, 1345-1623, 1348-1770, 1352-1619, 1353-1624, 1384-1750, 1384-1997, 1392-1661, 1439-1709, 1558-1769, 1558-1938, 1614-1899, 1725-2254, 1801-2046, 1859-2113, 1951-2195, 1951-2505, 1951-2603, 1953-2098, 1954-2495, 1956-2480, 1959-2116, 1963-2501, 1963-2511, 1967-2209, 1980-2512, 1988-2601, 2010-2582, 2010-2613, 2046-2595, 2063-2584, 2064-2588, 2079-2616, 2085-2358, 2092-2257, 2098-2617, 2103-2502, 2106-2616, 2114-2616, 2122-2407, 2128-2592, 2138-2592, 2141-2598, 2142-2595, 2142-2598, 2147-2599, 2148-2614, 2166-2616, 2167-2599, 2175-2307, 2182-2595, 2200-2595, 2201-2623, 2202-2595, 2207-2592, 2226-2597, 2229-2595, 2235-2505, 2244-2595, 2246-2616, 2263-2595, 2264-2542, 2269-2565, 2288-2595, 2288-2597, 2297-2592, 2324-2578, 2344-2595, 2372-2615, 2424-2615, 2535-2599

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
73/2159545CB1/ 4518	1-625, 1-695, 9-773, 46-2189, 275-1051, 343-822, 704-4518, 987-1569, 990-1563, 992-1230, 992-1443, 1029-1388, 1260-1989, 1262-1572, 1263-1572, 1266-1569, 1266-1572, 1314-1570, 1363-1613, 1363-1912, 1520-2058, 1588-2056, 1588-2086, 1590-1883, 1660-2224, 1736-2117, 1747-1939, 1781-2041, 1781-2324, 1798-2058, 1912-2216
74/8560269CB1/ 1238	1-263, 1-272, 1-329, 65-281, 65-332, 67-209, 90-274, 90-389, 91-270, 132-788, 331-583, 346-802, 348-620, 358-949, 509-1075, 559-1024, 619-1030, 620-826, 620-879, 620-1031, 666-875, 666-1129, 678-1093, 736-1091, 764-1057, 783-1107, 785-1045, 785-1078, 785-1088, 918-1184, 1020-1238
75/8710302CB1/ 1771	1-226, 82-368, 96-381, 125-423, 180-915, 184-395, 188-437, 189-379, 481-984, 493-989, 499-917, 500-980, 550-869, 558-991, 561-973, 561-991, 563-991, 565-1159, 569-1250, 606-1103, 667-1226, 688-1356, 714-974, 714-1216, 718-989, 771-1028, 844-1142, 846-1402, 914-1493, 943-1371, 1032-1571, 1041-1598, 1041-1623, 1091-1369, 1091-1372, 1091-1700, 1110-1383, 1154-1723, 1158-1444, 1273-1771, 1327-1598, 1330-1616
76/6778214CB1/ 2909	1-143, 60-830, 148-772, 165-441, 182-425, 207-373, 217-893, 342-479, 381-851, 400-885, 404-851, 407-764, 415-886, 420-851, 431-851, 432-867, 438-1047, 455-851, 475-882, 501-851, 515-851, 541-851, 550-813, 551-859, 568-851, 580-844, 621-880, 694-875, 871-1537, 882-1105, 883-1469, 884-1313, 888-1480, 891-1318, 891-1326, 891-1348, 891-1575, 911-1111, 911-1312, 914-1221, 925-1362, 925-1394, 1001-1290, 1066-1374, 1088-1243, 1113-1296, 1171-1645, 1213-1326, 1319-1556, 1319-1878, 1319-1919, 1351-1921, 1357-1592, 1376-1867, 1390-1921, 1406-1646, 1430-1671, 1478-1725, 1502-1701, 1522-1921, 1558-1917, 1560-1746, 1560-1921, 1587-1795, 1695-1921, 1841-2143, 1847-1921, 1922-2129, 1922-2425, 2000-2203, 2009-2249, 2016-2218, 2016-2313, 2019-2155, 2055-2290, 2096-2381, 2166-2394, 2204-2485, 2204-2486, 2204-2795, 2208-2440, 2208-2444, 2241-2443, 2246-2488, 2246-2831, 2286-2909

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
77/258383CB1/ 2216	1-226, 82-368, 82-513, 82-520, 82-540, 82-557, 82-571, 82-575, 82-610, 82-613, 82-619, 82-626, 82-628, 82-676, 82-734, 83-468, 83-571, 83-634, 88-548, 90-219, 96-381, 98-616, 122-423, 125-562, 168-785, 177-846, 182-513, 184-395, 188-437, 188-723, 188-748, 189-379, 189-663, 189-768, 201-747, 202-1007, 221-578, 234-785, 267-708, 354-910, 388-1073, 397-978, 411-979, 423-964, 437-928, 438-1036, 471-807, 480-1043, 529-962, 537-1184, 560-1084, 574-1077, 592-1010, 593-1073, 651-1084, 654-1066, 654-1084, 656-1084, 658-1252, 662-1343, 699-1196, 760-1319, 781-1449, 807-1067, 807-1309, 811-1082, 864-1121, 937-1235, 939-1495, 1007-1586, 1036-1464, 1072-1356, 1110-1417, 1125-1664, 1134-1691, 1134-1716, 1184-1462, 1184-1465, 1184-1793, 1203-1476, 1233-1424, 1247-1816, 1251-1537, 1366-1929, 1420-1691, 1423-1709, 1595-2216, 1643-2062, 1645-2216, 1669-2216, 1735-2216
78/2804937CB1/ 5320	1-841, 4-671, 292-843, 292-847, 294-643, 329-4681, 436-801, 605-894, 1232-1669, 1471-1741, 1950-2226, 1950-2439, 1950-2443, 1950-2520, 1950-2524, 2113-2740, 2115-2775, 2120-2793, 2213-2798, 2260-2857, 2323-2857, 2347-2870, 2383-3024, 2398-3079, 2429-3142, 2456-3035, 2459-3079, 2508-2838, 2532-3079, 2572-3160, 2599-3079, 2607-3209, 2607-3232, 2620-3219, 2697-3259, 2701-2879, 2701-2994, 2701-3340, 2763-2817, 2779-3422, 2804-3070, 2811-3035, 2831-3243, 2831-3417, 2844-3133, 2861-3369, 2900-3064, 2914-3500, 2974-3256, 3001-3609, 3021-3595, 3032-3535, 3060-3333, 3060-3553, 3065-3337, 3065-3565, 3085-3670, 3087-3678, 3094-3607, 3118-3607, 3118-3727, 3119-3762, 3129-3619, 3130-3693, 3135-3726, 3165-3841, 3166-3623, 3166-3766, 3214-3444, 3219-3809, 3223-3819, 3235-3715, 3240-3879, 3264-3710, 3305-3845, 3306-3820, 3329-3979, 3336-3792, 3339-3972, 3342-4018, 3369-4032, 3370-3890, 3405-3946, 3422-3639, 3432-3927, 3447-3706, 3471-4052, 3471-4124, 3493-4138, 3494-4126, 3496-4105, 3501-3762, 3501-3921, 3502-4028, 3503-4076, 3523-4138, 3556-3829, 3557-3772, 3569-4175, 3578-4188, 3591-3850, 3617-4269, 3621-3827, 3626-4175, 3637-3955, 3643-4208,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
78 (cont.)	3647-4203, 3659-4214, 3663-4155, 3663-4275, 3678-4011, 3682-3950, 3682-4088, 3682-4214, 3695-4268, 3701-4230, 3704-4293, 3717-3955, 3717-4286, 3726-4266, 3731-4312, 3734-4345, 3737-4355, 3749-4344, 3750-4006, 3750-4331, 3751-4011, 3756-4343, 3763-4293, 3780-4272, 3780-4450, 3781-4316, 3786-4056, 3807-4196, 3808-3905, 3809-4469, 3810-4368, 3813-4209, 3816-4344, 3821-4013, 3824-4305, 3831-4422, 3832-4360, 3848-4565, 3855-4565, 3857-4401, 3863-4151, 3863-4216, 3863-4323, 3868-4458, 3870-4509, 3871-4565, 3883-4415, 3886-4495, 3908-4535, 3914-4470, 3915-4384, 3924-4507, 3925-4182, 3934-4236, 3935-4427, 3937-4200, 3942-4356, 3942-4479, 3943-4218, 3946-4232, 3947-4591, 3948-4462, 3948-4591, 3952-4196, 3955-4129, 3961-4547, 3963-4514, 3967-4606, 3974-4608, 3978-4440, 3994-4458, 3997-4385, 3997-4619, 3998-4458, 4007-4470, 4009-4517, 4014-4509, 4017-4625, 4017-4682, 4021-4449, 4029-4599, 4039-4646, 4061-4094, 4062-4609, 4067-4366, 4074-4584, 4076-4659, 4081-4349, 4086-4313, 4086-4589, 4097-4659, 4120-4551, 4122-4741, 4124-4716, 4125-4591, 4135-4670, 4136-4645, 4137-4683, 4142-4709, 4146-4372, 4146-4634, 4146-4692, 4146-4729, 4147-4396, 4153-
	4715, 4162-4679, 4164-4706, 4169-4639, 4169-4699, 4172-4686, 4180-4403, 4186-4718, 4188-4810, 4192-4842, 4211-4753, 4230-4681, 4244-4672, 4245-4836, 4258-4729, 4263-4622, 4263-4854, 4264-4699, 4275-4830, 4276-4885, 4280-4888, 4288-4534, 4288-4723, 4288-4938, 4289-4895, 4297-4429, 4300-4871, 4312-4657, 4315-4760, 4315-4914, 4327-4753, 4327-4806, 4338-4888, 4340-4761, 4340-4903, 4355-4590, 4357-4565, 4361-4899, 4364-5013, 4374-4899, 4389-4894, 4389-4898, 4390-4903, 4393-4964, 4399-4648, 4399-4895, 4406-4719, 4416-4654, 4416-4664, 4416-4879, 4421-4899, 4434-4902, 4438-4903, 4448-4978, 4471-4903, 4475-5125, 4477-4836, 4494-4853, 4494-4958, 4496-4903, 4502-4901, 4503-4914, 4507-4903, 4510-4902, 4568-4974, 4570-4762, 4573-4903, 4575-4805, 4614-4898, 4651-4903, 4720-4943, 4720-4964, 4722-4893, 4867-5130, 4953-5320, 5018-5269, 5104-5320, 5134-5320, 5155-5320
79/7494915CB1/ 653	1-653, 51-602
80/2073751CB1/ 1794	1-415, 14-312, 14-367, 14-496, 14-507, 14-510, 14-559, 14-560, 19-268, 19-463, 54-340, 62-542, 81-564, 102-560, 110-941, 111-372, 117-369, 117-599, 120-326, 191-508, 199-709, 239-550, 252-731, 532-599, 600-658, 788-1715, 790-1227, 792-1044, 1051-1229, 1082-1227, 1266-1794, 1280-1583, 1395-1544, 1716-1794

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
81/3178841CB1/ 2691	1-282, 1-510, 8-281, 53-404, 55-518, 55-519, 58-501, 73-519, 78-447, 86-348, 90-501, 97-886, 115-515, 120-324, 120-446, 120-519, 130-519, 136-387, 136-694, 148-519, 152-519, 160-515, 163-517, 163-519, 168-466, 179-464, 202-510, 265-954, 267-469, 467-1036, 485-1029, 512-970, 514-1021, 515-1164, 516-907, 516-981, 517-1116, 518-1067, 521-905, 521-906, 521-931, 521-946, 521-991, 521-997, 521-1021, 521-1022, 521-1023, 521-1036, 521-1041, 521-1047, 521-1050, 521-1059, 521-1066, 521-1082, 521-1089, 521-1113, 521-1115, 521-1126, 521-1128, 521-1130, 521-1136, 521-1145, 521-1154, 521-1158, 521-1167, 521-1172, 522-1116, 527-1206, 530-871, 530-1209, 543-815, 543-921, 547-867, 547-977, 559-1076, 561-929, 585-1217, 599-882, 613-1267, 619-885, 620-1194, 647-1261, 653-1303, 673-1310, 692-1369, 698-1123, 720-1378, 733-1476, 734-889, 738-1379, 744-1025, 764-1217, 770-1470, 789-985, 793-1470, 794-1304, 800-1055, 813-1416, 834-1496, 871-1568, 883-1152, 883-1163, 883-1649, 905-1471, 925-1475, 937-1203, 1011-1630, 1012-1589, 1031-1512, 1054-1639, 1066-1719, 1120-1606, 1120-1755, 1156-1396, 1157-1755, 1159-1431, 1177-1755, 1184-1388, 1238-1733, 1247-1737, 1248-1805,
	1273-1512, 1273-1519, 1273-1527, 1273-1528, 1303-1769, 1346-1906, 1357-1525, 1361-1755, 1430-1727, 1453-1966, 1459-1762, 1466-1755, 1492-2116, 1499-1784, 1501-1805, 1544-2110, 1565-1947, 1580-1967, 1582-1816, 1588-1765, 1605-1771, 1622-2150, 1631-2200, 1655-2263, 1672-1961, 1672-2161, 1717-2229, 1718-1921, 1718-2180, 1730-2225, 1749-2302, 1749-2382, 1751-2005, 1751-2007, 1766-1992, 1776-2037, 1803-2200, 1822-2083, 1856-2292, 1869-2440, 1918-2081, 1939-2549, 1949-2601, 1953-2601, 1968-2397, 1989-2574, 2004-2394, 2009-2536, 2043-2332, 2067-2346, 2068-2333, 2068-2425, 2069-2318, 2069-2417, 2074-2327, 2083-2341, 2112-2378, 2112-2648, 2112-2687, 2132-2442, 2136-2682, 2151-2662, 2167-2380, 2187-2691, 2200-2462, 2207-2486, 2234-2531, 2238-2467, 2417-2579
82/3674807CB1/ 2056	1-290, 1-436, 1-453, 1-456, 1-488, 1-491, 1-524, 1-558, 1-563, 1-581, 1-591, 1-598, 1-600, 1-613, 1-641, 1-652, 1-686, 1-754, 2-575, 41-736, 58-502, 83-748, 105-854, 126-911, 133-395, 157-874, 167-650, 197-807, 226-703, 230-807, 265-716, 323-891, 388-1071, 406-708, 417-1080, 440-1053, 441-682, 443-1099, 444-1024, 445-1075, 497-850, 504-1064, 527-1165, 547-1182, 573-1124, 631-1179, 635-1204, 636-1165, 660-1163, 664-1218, 688-864, 716-1206, 719-1339, 723-1023, 752-1255, 759-1295, 771-1325, 773-1382, 778-1378, 795-1382, 800-1455, 804-1354, 811-1272, 823-1461, 834-1202, 837-1525, 845-1438, 846-1296, 850-1224, 857-1058, 866-983, 866-1282, 866-1362, 918-1530, 976-1553, 983-1497, 1009-1446, 1046-1509, 1251-1882, 1388-1888, 1424-1711, 1721-2056, 1754-2056, 1807-2056

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
83/1794922CB1/ 2275	1-283, 1-407, 169-592, 169-957, 579-972, 838-1488, 850-1634, 896-1138, 896-1385, 1031-1642, 1084-1347, 1084-1544, 1166-1411, 1357-1649, 1357-1662, 1417-2118, 1459-1655, 1459-1722, 1464-1719, 1464-1996, 1472-2167, 1574-2254, 1575-1878, 1643-1791, 1684-2275, 1739-1994, 1739-2183, 1739-2199, 1765-2246, 1808-2266, 1810-2269, 1816-2265, 1821-2257, 1825-2275, 1844-2265, 1876-2266, 1926-2265, 2042-2275
84/1795509CB1/ 1219	1-282, 1-284, 62-313, 71-271, 72-325, 72-512, 72-826, 100-415, 108-345, 145-648, 167-555, 238-443, 238-943, 240-788, 293-552, 322-845, 337-887, 370-887, 397-685, 443-683, 443-921, 443-943, 469-694, 499-1096, 512-745, 526-717, 530-845, 571-1009, 574-1034, 575-889, 579-1038, 579-1051, 579-1061, 583-1052, 584-845, 651-878, 661-1210, 661-1219, 668-1219, 674-1190, 677-845, 740-1208, 859-1219, 944-1206
85/2017180CB1/ 1015	1-264, 1-305, 1-447, 6-433, 7-515, 7-517, 9-330, 11-615, 15-308, 16-235, 16-293, 18-211, 18-279, 18-295, 20-665, 21-224, 22-375, 26-718, 33-259, 35-300, 35-812, 36-290, 39-504, 40-334, 51-328, 52-287, 52-747, 58-317, 72-686, 79-527, 80-344, 114-677, 145-308, 174-444, 196-465, 228-463, 252-753, 362-768, 439-717, 459-1015, 479-748, 520-657, 527-1012, 557-862, 602-884, 673-1012, 703-994, 706-987, 722-964, 725-976, 725-1015, 731-999, 743-991
86/219442CB1/ 2392	1-408, 1-416, 1-460, 2-384, 4-297, 7-283, 118-745, 158-447, 350-855, 376-855, 381-612, 394-664, 419-855, 441-855, 493-659, 525-1114, 592-855, 629-1263, 635-953, 665-1113, 774-1133, 801-998, 820-1269, 883-1239, 913-1150, 916-1044, 937-1114, 963-1061, 963-1115, 1035-1269, 1096-1628, 1120-1616, 1437-1941, 1617-2266, 1619-1890, 1668-1976, 1691-2045, 1707-2253, 1713-2392, 1787-2080, 1798-2096, 1816-1946, 1818-2392, 1836-2392, 1879-2345, 1901-2392, 1935-2328, 1939-2392, 1944-2328, 1953-2059, 1964-2382
87/2597459CB1/ 1799	1-578, 148-611, 404-1009, 421-824, 427-884, 432-612, 517-1000, 643-847, 643-909, 643-1098, 643-1123, 643-1130, 643-1136, 643-1139, 643-1143, 643-1149, 643-1150, 643-1164, 643-1169, 643-1178, 643-1180, 643-1181, 643-1182, 643-1189, 643-1200, 643-1210, 643-1219, 643-1249, 643-1277, 643-1290, 674-1283, 701-1279, 705-1413, 755-1305, 898-1080, 966-1231, 966-1425, 1018-1711, 1019-1588, 1022-1409, 1052-1630, 1066-1552, 1069-1442, 1119-1713, 1133-1696, 1145-1555, 1165-1754, 1190-1795, 1194-1767, 1202-1743, 1251-1761, 1255-1381, 1262-1789, 1262-1799, 1283-1799, 1314-1440, 1314-1767, 1379-1634, 1384-1599, 1384-1799, 1437-1706, 1444-1727

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
88/2783863CB1/ 3718	1-556, 227-772, 270-767, 279-574, 279-823, 279-909, 300-863, 387-977, 468-959, 600-1232, 884-1201, 884-1341, 884-1455, 927-1069, 962-1117, 962-1562, 985-1742, 1007-1525, 1007-1592, 1045-1763, 1066-1448, 1088-1339, 1171-2106, 1183-1451, 1183-1742, 1189-1967, 1250-1897, 1267-1954, 1296-2124, 1301-1449, 1433-2059, 1466-2059, 1560-2297, 1571-1823, 1571-1828, 1606-2363, 1625-1911, 1658-2015, 1664-1905, 1664-1919, 1664-2146, 1676-2352, 1681-1994, 1683-2031, 1686-2246, 1839-2104, 1846-2049, 1850-2069, 1856-2826, 1870-2666, 2056-2233, 2113-2365, 2216-2491, 2218-2823, 2241-2473, 2256-2834, 2256-2881, 2266-2546, 2298-2564, 2330-2618, 2336-2618, 2341-2784, 2363-2604, 2377-2897, 2387-2902, 2402-2646, 2402-2671, 2402-2954, 2450-2733, 2456-2875, 2503-2721, 2512-2973, 2627-2889, 2644-2908, 2741-2973, 2755-3289, 2816-3209, 2836-3387, 2847-3655, 2895-3151, 2896-3466, 2912-3158, 2945-3180, 2970-3262, 2971-3224, 2985-3710, 2987-3274, 3063-3718, 3064-3675, 3106-3636, 3116-3710, 3117-3695, 3139-3704, 3143-3712, 3192-3700, 3200-3496, 3201-3712, 3215-3664, 3227-3699, 3228-3645, 3248-3711, 3249-3711, 3252-3710, 3258-3710, 3305-3665, 3356-3599, 3356-3696, 3356-3710, 3372-3707, 3407-3594, 3407-3605, 3407-3707, 3413-3664, 3426-3711, 3429-3707
89/2902971CB1/ 3250	1-105, 1-622, 1-644, 2-200, 2-285, 2-287, 2-821, 4-273, 8-232, 8-254, 8-256, 8-282, 8-635, 13-251, 16-236, 18-138, 18-284, 18-303, 18-313, 19-812, 29-673, 34-231, 34-232, 34-282, 34-842, 43-482, 61-274, 64-266, 65-266, 68-266, 102-650, 200-274, 208-258, 276-582, 327-501, 575-1157, 665-838, 665-1384, 784-932, 803-1417, 848-1665, 855-1472, 908-1530, 1013-1615, 1030-1345, 1041-1266, 1123-1527, 1157-1611, 1160-1737, 1293-1535, 1403-1721, 1450-2069, 1466-1710, 1631-1749, 1631-2071, 1700-1983, 1739-2028, 1739-2180, 1809-2254, 1844-2278, 1897-2310, 1959-2199, 1959-2477, 1959-2491, 1971-2250, 2001-2477, 2043-2477, 2043-2516, 2050-2477, 2063-2477, 2165-2362, 2165-2603, 2165-2783, 2285-2980, 2363-2748, 2364-2975, 2394-2659, 2467-3133, 2477-3163, 2478-2936, 2478-2978, 2536-2807, 2604-2987, 2623-2875, 2623-2907, 2623-2911, 2623-2924, 2625-2944, 2678-3224, 2698-3219, 2755-3204, 2766-3237, 2775-3250, 2789-3062, 2789-3217, 2789-3222, 2792-3245, 2801-3242, 2812-3237, 2828-3237, 2845-3235, 2849-3228, 2897-3231, 2923-3131, 2953-3250
90/368660CB1/ 2295	1-584, 4-273, 7-286, 33-410, 34-790, 35-811, 41-548, 529-1215, 530-935, 584-1232, 648-1227, 653-1394, 654-1118, 684-1187, 795-1048, 923-1158, 928-1227, 928-1415, 958-1196, 1016-1265, 1017-1325, 1038-1309, 1163-1368, 1172-1542, 1388-1705, 1473-1557, 1477-1691, 1482-1774, 1485-2117, 1489-1747, 1526-2130, 1569-1748, 1658-2186, 1712-1917, 1712-1931, 1712-1934, 1712-1936, 1712-1947, 1712-1965, 1712-1982, 1714-2190, 1720-1853, 1728-2295, 1740-1996, 1754-2252, 1777-1984, 1779-2042, 1791-2051, 1853-2031, 1863-1997, 1920-2293, 1922-2295

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
91/2804990CB1/ 1777	1-738, 1-780, 210-758, 214-1734, 414-722, 484-517, 495-743, 498-884, 518-801, 641-1240, 679-975, 735-1131, 743-1234, 763-1285, 847-1458, 868-1191, 918-1186, 973-1353, 1041-1348, 1041-1452, 1094-1432, 1154-1594, 1202-1395, 1214-1614, 1214-1718, 1218-1688, 1218-1710, 1218-1747, 1222-1487, 1222-1689, 1222-1746, 1253-1478, 1255-1777, 1266-1720, 1270-1594, 1271-1594, 1273-1594, 1279-1550, 1282-1552
92/168571CB1/ 3181	1-564, 1-591, 50-722, 177-780, 382-983, 404-661, 426-1056, 444-979, 543-1149, 599-1147, 648-1148, 682-1146, 687-1146, 689-1139, 909-1605, 1150-1412, 1150-1567, 1150-1611, 1152-1417, 1157-1765, 1159-1624, 1159-1666, 1159-1759, 1159-1760, 1168-1625, 1196-1449, 1216-1550, 1291-1505, 1291-1541, 1322-1737, 1331-1542, 1579-1793, 1601-1967, 1707-2028, 1717-1942, 1781-2282, 1800-2145, 1836-2063, 1839-2077, 1839-2409, 1865-2430, 1875-2369, 1952-2369, 1960-2574, 1960-2593, 1961-2169, 1968-2611, 1971-2361, 1975-2558, 1979-2167, 1999-2656, 2019-2298, 2021-2406, 2021-2476, 2022-2241, 2022-2559, 2062-2304, 2079-2271, 2112-2382, 2128-2715, 2133-2685, 2160-2672, 2185-2509, 2204-2808, 2222-2644, 2223-2814, 2293-2571, 2303-2523, 2305-2868, 2329-2576, 2334-2606, 2335-2628, 2341-2856, 2379-2693, 2417-2638, 2425-2727, 2472-2896, 2494-3097, 2495-2724, 2500-2696, 2500-2772, 2500-2781, 2500-2835, 2500-3078, 2501-2773, 2504-3008, 2519-3171, 2540-2851, 2544-2861, 2565-2827, 2567-2786, 2579-3101, 2583-2744, 2586-2885, 2586-3094, 2588-3102, 2602-2863, 2607-2859, 2648-2863, 2649-2814, 2664-3165, 2671-3161, 2708-3002, 2749-3094, 2800-3180, 2821-3098, 2821-3116, 2902-3149, 3025-3181
93/1286391CB1/ 5987	1-650, 1-2676, 152-347, 266-860, 266-968, 362-528, 442-4155, 634-1105, 660-901, 823-1005, 823-1263, 1004-1581, 1104-1337, 1126-1735, 1207-1630, 1229-1926, 1324-2013, 1415-1659, 1415-1728, 1838-2104, 1872-2353, 1959-2529, 2093-2721, 2111-2647, 2132-2438, 2177-2678, 2233-2705, 2331-3122, 2352-2542, 2447-2678, 2491-3289, 2537-3129, 2704-3268, 2729-3009, 2729-3192, 2729-3194, 2729-3279, 2729-3292, 2805-3377, 2822-3285, 2875-3573, 2893-3549, 3028-3710, 3060-3611, 3079-3290, 3114-3527, 3256-3902, 3281-3856, 3297-3848, 3331-3838, 3336-3851, 3364-3863, 3438-3903, 3439-4143, 3545-4169, 3603-4056, 3603-4252, 3698-4456, 3701-4276, 3742-4452, 3743-4167, 3743-4326, 3743-4360, 3743-4365, 3743-4411, 3743-4424, 3743-4427, 3743-4458, 3743-4487, 3747-3967, 3763-4251, 3780-4419, 3790-4053, 3790-4220, 3804-4498, 3807-4424, 3887-4391, 3888-4502, 3894-4451, 3896-4502, 3928-4074, 3928-4156, 3928-4499, 3930-4380, 3933-4609, 3961-4662, 3986-4244, 3991-4459, 3992-4629, 3999-4529, 4025-4577, 4054-4691, 4054-4709, 4145-4184, 4193-4757, 4350-4878, 4350-5055, 4350-5067, 4351-5002, 4490-5197, 4490-5265, 4532-5209, 4545-5117, 4563-5157, 4592-5083, 4646-5278,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
93 (cont.)	4650-5278, 4652-5315, 4683-5357, 4684-5339, 4709-5193, 4715-5341, 4720-5399, 4731-5400, 4732-5395, 4744-5220, 4756-5348, 4763-5378, 4785-5413, 4788-5078, 4788-5267, 4788-5281, 4788-5309, 4788-5383, 4788-5472, 4788-5491, 4789-5386, 4789-5395, 4791-5425, 4792-5440, 4798-5064, 4798-5363, 4799-5383, 4799-5491, 4799-5584, 4800-5349, 4801-5607, 4818-5087, 4823-5300, 4831-5288, 4844-5125, 4844-5320, 4845-5032, 4861-5409, 4897-5514, 4898-5493, 4899-5527, 4906-5384, 4910-5163, 4910-5578, 4914-5480, 4914-5493, 4915-5158, 4915-5503, 4920-5527, 4921-5496, 4936-5650, 4939-5221, 4942-5449, 4946-5535, 4949-5559, 4950-5527, 4954-5455, 4954-5536, 4956-5501, 4964-5597, 4965-5565, 4967-5599, 4980-5564, 4999-5483, 5001-5457, 5002-5527, 5018-5556, 5021-5274, 5024-5534, 5040-5564, 5042-5514, 5061-5522, 5068-5539, 5070-5696, 5140-5371, 5232-5942, 5311-5977, 5327-5981, 5331-5987, 5333-5977, 5336-5955, 5485-5987, 5572-5981, 5629-5743, 5629-5920, 5684-5905, 5688-5920, 5753-5965, 5766-5920
94/2007684CB1/ 765	1-275, 1-536, 23-227, 55-580, 250-765, 345-610, 420-765, 451-759, 452-613, 474-613
95/2227040CB1/ 2674	1-418, 123-418, 162-484, 167-671, 187-391, 428-651, 524-764, 524-769, 524-1060, 552-667, 601-952, 797-1050, 797-1336, 840-1091, 1023-1444, 1034-1555, 1104-1684, 1149-1369, 1149-1626, 1301-1416, 1310-1545, 1318-1546, 1506-2099, 1661-2031, 1662-1944, 1677-2331, 1705-1968, 1771-2277, 1799-2106, 1901-2385, 1936-2142, 1936-2190, 1936-2208, 1941-2209, 1944-2242, 1944-2247, 1944-2400, 1945-2150, 1968-2192, 1968-2474, 2002-2233, 2035-2674, 2039-2295, 2064-2401, 2071-2639, 2083-2256, 2097-2248, 2101-2671, 2160-2419, 2166-2330, 2166-2377, 2168-2389, 2278-2377, 2290-2377
96/4346130CB1/ 5920	1-479, 2-493, 2-533, 2-535, 2-561, 21-272, 24-400, 30-558, 32-402, 37-480, 179-705, 471-800, 492-789, 536-672, 536-910, 536-1141, 784-1423, 860-1379, 971-1215, 971-1226, 971-1358, 971-1565, 1053-1324, 1137-1485, 1349-1817, 1368-1624, 1368-1841, 1435-1987, 1435-2009, 1561-2121, 1588-1817, 1590-2121, 1600-1817, 1720-1817, 1722-1817, 1768-2170, 2013-2504, 2079-2317, 2079-2346, 2079-2398, 2155-2682, 2155-2738, 2310-5524, 2373-2857, 2451-2684, 2544-3072, 2771-3257, 2948-3187, 3369-4024, 3537-4052, 4366-4776, 4377-4732, 4378-4682, 4413-4693, 4752-5051, 4919-5161, 4945-5574, 5013-5385, 5055-5646, 5070-5319, 5241-5501, 5303-5864, 5392-5780, 5392-5802, 5393-5627, 5431-5684, 5431-5920, 5433-5837, 5469-5873
97/55117040CB1/ 1689	1-341, 1-347, 1-385, 1-418, 1-560, 1-661, 4-661, 5-661, 58-637, 89-157, 89-167, 89-171, 114-823, 200-268, 200-282, 201-282, 225-268, 325-877, 805-1689, 817-1532

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
98/7472392CB1/ 1314	1-157, 1-158, 1-195, 1-333, 1-364, 1-438, 1-508, 1-510, 1-513, 1-546, 1-620, 1-624, 1-657, 1-684, 1-721, 1-735, 1-744, 1-764, 1-831, 1-912, 1-1048, 1-1307, 4-548, 14-811, 20-661, 24-920, 76-1044, 80-141, 231-778, 231-782, 342-761, 502-1031, 536-1035, 536-1314, 561-657, 576-973, 648-788
99/4028960CB1/ 3322	1-824, 11-856, 19-545, 97-209, 115-790, 174-777, 174-850, 187-655, 209-534, 396-678, 434-650, 434-950, 657-935, 657-1053, 783-1046, 783-1383, 834-970, 879-1067, 885-1496, 945-1592, 1071-1491, 1176-1337, 1234-1892, 1284-1663, 1293-1801, 1314-1534, 1320-1704, 1341-1953, 1454-1760, 1684-1919, 1684-1920, 1684-1933, 1684-2140, 1690-2148, 1850-2042, 1873-2116, 1875-2098, 1890-2348, 1963-2703, 2027-2518, 2043-2329, 2101-2324, 2127-2372, 2137-2776, 2169-2577, 2218-2493, 2322-2911, 2324-2491, 2336-2994, 2344-2558, 2395-2924, 2403-3035, 2409-2592, 2414-2978, 2415-2688, 2496-2757, 2498-2724, 2505-3038, 2529-2864, 2536-2980, 2562-2692, 2632-3052, 2654-3051, 2743-3233, 2767-3227, 2784-3224, 2787-2921, 2794-3016, 2808-3218, 2817-3050, 2906-3322
100/8227004CB1/ 3621	1-250, 1-434, 1-492, 1-611, 1-612, 1-676, 1-860, 100-900, 114-772, 238-680, 247-879, 273-494, 288-685, 341-1108, 396-813, 402-813, 425-1104, 467-1245, 507-1224, 526-896, 650-1364, 655-1231, 656-1302, 661-1301, 668-1423, 671-1042, 679-1258, 711-1301, 760-1301, 763-1131, 782-1542, 787-1323, 794-1224, 818-1271, 852-1355, 874-1462, 904-1036, 916-1301, 961-1624, 980-1454, 1011-1397, 1025-1592, 1028-1255, 1031-1728, 1065-1298, 1065-1605, 1149-1842, 1168-1548, 1202-1682, 1209-2160, 1217-1697, 1229-1757, 1282-1511, 1313-1804, 1334-2037, 1362-2087, 1372-1967, 1400-2375, 1407-2144, 1497-2162, 1621-2285, 1627-2092, 1627-2569, 1629-2068, 1630-2318, 1652-1880, 1673-2409, 1683-2354, 1708-1979, 1733-1995, 1742-2449, 1743-2000, 1745-2189, 1746-2419, 1751-2439, 1751-2497, 1759-2530, 1792-2108, 1795-2614, 1816-2097, 1834-2105, 1845-2598, 1893-2144, 1893-2485, 1903-2586, 1904-2715, 1904-2732, 1949-2203, 1951-2207, 1975-2296, 1977-2878, 1979-2214, 1980-2645, 1989-2407, 1996-2256, 1999-2668, 2006-2722, 2025-2560, 2032-2514, 2046-2745, 2061-2540, 2079-2707, 2094-2348, 2094-2895, 2106-2296, 2106-2510, 2109-2289, 2109-2597, 2112-2372, 2119-2385, 2138-2789,
	2150-2447, 2174-2333, 2188-2889, 2195-2841, 2209-2802, 2216-2546, 2251-2596, 2251-2604, 2272-2527, 2296-2583, 2297-2575, 2312-2811, 2321-2574, 2335-2664, 2343-2950, 2350-2865, 2351-2879, 2353-2827, 2355-2567, 2363-2759, 2390-2588, 2394-2664, 2421-2920, 2426-2740, 2427-2942, 2447-2862, 2458-2846, 2462-3141, 2466-2720, 2470-3192, 2483-2699, 2483-2997, 2798-3256, 2859-3106, 2908-3175, 2921-3178, 3049-3309, 3049-3621

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
101/3044763CB1/ 2704	1-283, 41-819, 66-368, 88-204, 215-856, 218-734, 303-647, 389-661, 421-706, 465-713, 482-1060, 490-610, 490-952, 511-1108, 516-1026, 569-1199, 576-1240, 585-1112, 673-1246, 677-1221, 697-974, 732-978, 742-1290, 755-1049, 780-1430, 824-997, 844-1079, 872-1392, 913-1188, 946-1561, 979-1529, 1035-1304, 1035-1755, 1035-1818, 1039-1294, 1056-1207, 1056-1290, 1076-1758, 1111-1489, 1113-1375, 1117-1433, 1147-1515, 1162-1458, 1166-1744, 1171-1427, 1172-1518, 1173-1393, 1191-1461, 1208-1511, 1217-1365, 1227-1537, 1249-1800, 1273-1923, 1290-1513, 1296-1547, 1302-1495, 1321-1409, 1324-1798, 1332-1627, 1341-1494, 1341-1623, 1343-1818, 1363-1682, 1363-1886, 1380-1649, 1388-1631, 1390-1635, 1390-1822, 1392-1911, 1397-1818, 1400-1818, 1402-1734, 1419-1791, 1420-1799, 1429-1818, 1438-1714, 1460-1646, 1461-1677, 1503-1918, 1511-1897, 1522-1608, 1556-1799, 1571-1818, 1577-2418, 1596-1818, 1637-1918, 1643-1878, 1658-1808, 1658-1818, 1660-2116, 1754-1860, 1757-1813, 1757-1818, 1760-1918, 2110-2704
102/4044519CB1/ 1345	1-146, 1-291, 1-490, 1-574, 1-583, 1-598, 1-627, 1-631, 1-657, 1-679, 1-680, 1-808, 92-804, 98-820, 159-495, 479-1252, 509-1289, 615-1345, 674-1329, 684-1341, 764-1342
103/71351918CB1/ 1607	1-246, 27-253, 29-80, 30-304, 33-278, 33-376, 33-561, 34-225, 36-571, 38-280, 43-324, 60-315, 61-286, 169-738, 308-578, 419-947, 423-659, 424-695, 441-1112, 444-972, 466-994, 468-793, 478-829, 490-798, 498-1124, 506-725, 512-1016, 516-946, 519-791, 545-1040, 547-997, 565-1145, 568-849, 570-813, 570-956, 571-965, 576-736, 579-1111, 581-1028, 582-1167, 587-1142, 588-1208, 595-1145, 595-1378, 596-857, 600-1109, 614-1186, 624-1220, 628-848, 633-1213, 638-1225, 640-1172, 641-1244, 648-822, 648-1403, 649-955, 659-756, 672-1106, 676-1403, 678-949, 680-1192, 683-931, 684-1206, 684-1228, 693-976, 699-1040, 700-960, 711-1305, 715-1263, 726-1304, 730-987, 730-1358, 742-1339, 748-1399, 757-1264, 760-1244, 770-992, 771-1131, 787-1301, 788-1014, 796-1301, 798-1336, 814-1374, 825-1536, 832-1068, 838-1396, 843-1044, 846-1093, 846-1482, 858-1114, 860-1494, 866-1149, 866-1428, 867-1306, 875-1471, 876-1120, 878-1146, 879-1238, 883-1528, 890-1533, 893-1090, 897-1393, 899-1127, 899-1546, 900-1203, 904-1281, 917-1364, 918-1430, 928-1371, 930-1544, 937-1540, 949-1444, 952-1414, 952-1537, 955-1527, 956-1081, 963-1533, 988-1106, 990-1264, 999-1540, 1003-1459, 1006-1539,
	1014-1473, 1015-1415, 1025-1288, 1028-1543, 1028-1550, 1030-1539, 1043-1527, 1046-1317, 1047-1358, 1047-1367, 1047-1607, 1050-1546, 1051-1475, 1058-1353, 1070-1316

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
104/8109363CB1/ 2622	1-521, 125-498, 246-543, 268-746, 315-549, 367-885, 403-1139, 464-717, 634-903, 662-1255, 704-925, 742-1371, 845-1124, 864-1385, 966-1471, 1041-1444, 1045-1444, 1105-1415, 1105-1618, 1320-1624, 1320-1888, 1361-1648, 1373-1884, 1448-1688, 1556-1811, 1584-2049, 1591-1840, 1596-2174, 1690-1962, 1697-1956, 1697-2207, 1701-1826, 1772-2051, 1790-2055, 1935-2190, 1937-2617, 2003-2253, 2003-2302, 2003-2490, 2010-2236, 2010-2309, 2011-2622, 2012-2593, 2052-2318, 2063-2292, 2063-2296, 2087-2622, 2094-2616, 2096-2340, 2161-2424, 2254-2510
105/1272746CB1/ 3489	1-214, 1-254, 1-595, 3-267, 7-248, 41-267, 46-275, 154-268, 202-394, 202-775, 202-841, 207-392, 207-534, 207-918, 340-632, 340-866, 401-1133, 463-1095, 506-1225, 592-864, 746-1357, 748-1382, 772-1095, 929-1193, 954-1410, 959-1425, 1059-1554, 1064-1421, 1166-1419, 1336-1669, 1449-1951, 1493-2128, 1519-1968, 1519-2004, 1519-2045, 1519-2055, 1519-2096, 1519-2113, 1519-2155, 1519-2161, 1519-2202, 1520-2009, 1521-1786, 1560-2144, 1569-2178, 1625-2247, 1642-1783, 1663-1912, 1700-2520, 1710-2140, 1734-2263, 1813-2446, 1860-2519, 1880-2572, 1937-2098, 1951-2552, 1953-2706, 1955-2186, 1955-2219, 1955-2641, 1955-2763, 1956-2325, 1959-2420, 1965-2559, 1976-2243, 1982-2480, 1994-2072, 2023-2579, 2024-2647, 2031-2642, 2052-2697, 2053-2742, 2126-2775, 2127-2699, 2135-2420, 2139-2833, 2173-2730, 2177-2840, 2182-2783, 2183-2869, 2215-2456, 2215-2830, 2229-2785, 2236-2935, 2242-2935, 2246-2852, 2256-3027, 2257-2763, 2288-2848, 2294-2924, 2296-2932, 2301-2525, 2314-2824, 2322-2524, 2322-2792, 2324-2935, 2341-2935, 2344-2914, 2354-2935, 2358-2935, 2373-2932, 2379-2932, 2382-2935, 2384-2918, 2386-2599, 2386-2935, 2389-2935, 2391-3001, 2397-2693, 2413-2639,
	2422-2935, 2424-2935, 2431-2679, 2443-3079, 2449-2935, 2468-3177, 2477-2941, 2487-2856, 2495-2935, 2499-2932, 2516-2932, 2524-3136, 2526-2932, 2530-2938, 2558-2935, 2561-2935, 2609-3294, 2620-3315, 2621-2940, 2624-2907, 2648-2773, 2658-3455, 2674-3489, 2708-3463, 2756-3191, 2768-2977, 2768-3252, 2768-3265, 2768-3267, 2768-3278, 2768-3288, 2768-3372, 2830-3068, 2830-3195, 2833-3452, 2867-3421, 2889-3431, 2916-3366, 2941-3196, 2962-3443, 2977-3197, 2992-3444, 3020-3285, 3105-3349, 3105-3434, 3105-3452, 3122-3330, 3125-3443, 3126-3449, 3137-3306, 3228-3443, 3230-3443, 3305-3449
106/1839974CB1/ 2269	1-747, 1-757, 24-624, 25-529, 25-530, 87-694, 150-944, 179-571, 313-528, 458-1022, 469-1197, 588-930, 820-1097, 893-1455, 1174-1547, 1174-1604, 1179-1519, 1362-1636, 1447-1743, 1513-1804, 1521-1764, 1654-2269, 1658-1892, 1658-2148, 1661-2231, 1812-2269, 1905-2103, 1912-2175, 1949-2240, 2008-2098

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
107/1877336CB1/ 3075	1-194, 1-197, 1-399, 3-412, 4-295, 7-222, 9-240, 13-569, 13-608, 15-558, 15-642, 17-218, 19-628, 25-277, 25-549, 33-305, 34-614, 38-621, 43-698, 102-510, 103-770, 103-800, 112-738, 137-754, 138-429, 144-664, 173-730, 241-514, 259-535, 262-614, 262-662, 286-594, 421-648, 573-1015, 641-1229, 691-1147, 713-942, 725-1421, 757-1368, 803-1368, 806-1309, 820-1492, 835-1360, 835-1368, 837-1459, 840-1325, 844-1339, 872-1439, 877-1464, 883-1467, 889-1366, 893-1557, 895-1366, 898-1370, 926-1593, 959-1455, 971-1387, 1007-1678, 1014-1567, 1045-1572, 1075-1707, 1097-1225, 1104-1471, 1104-1662, 1104-1698, 1106-1800, 1154-1653, 1156-1735, 1160-1850, 1176-1709, 1176-1859, 1177-1806, 1180-1801, 1183-1718, 1196-1800, 1215-1736, 1224-1878, 1252-1897, 1265-1829, 1272-1801, 1283-1639, 1291-1815, 1295-1830, 1314-1866, 1316-1807, 1318-1861, 1319-1795, 1319-1908, 1325-1836, 1335-1848, 1335-1851, 1338-1941, 1340-1905, 1346-1784, 1347-1867, 1355-1820, 1359-1851, 1363-1859, 1379-1947, 1384-1891, 1406-1927, 1417-1902, 1436-2104, 1500-2058, 1519-1981, 1520-2025, 1520-2034, 1521-1859, 1521-1860, 1526-2061, 1528-2160, 1529-2025, 1548-2212, 1557-2158, 1562-2158, 1567-2301,
	1591-2113, 1598-2183, 1606-2125, 1610-2247, 1612-2157, 1612-2186, 1628-2291, 1647-2255, 1675-2326, 1684-2069, 1684-2301, 1687-2349, 1699-2148, 1700-2215, 1700-2252, 1704-2212, 1705-2205, 1705-2360, 1738-2370, 1743-2076, 1750-1776, 1752-2278, 1762-2337, 1763-2190, 1763-2285, 1764-2236, 1777-2503, 1791-2295, 1793-2186, 1794-2364, 1800-2361, 1805-2226, 1815-2085, 1819-2382, 1831-2323, 1842-2302, 1873-2325, 1988-2212, 2005-2608, 2039-2706, 2085-2673, 2323-2357, 2365-2695, 2397-2962, 2402-2686, 2418-2677, 2419-2999, 2420-2670, 2430-2668, 2448-2678, 2449-2612, 2457-2660, 2603-2663, 2702-2789, 2704-3056, 2716-2984, 2723-3056, 2729-3057, 2736-2912, 2738-2981, 2762-2981, 2762-2997, 2769-2974, 2769-3013, 2769-3027, 2777-3057, 2784-3024, 2785-3058, 2789-3029, 2801-3018, 2811-3056, 2811-3075, 2822-3049, 2823-3049, 2829-3055, 2831-3075, 2833-3039, 2834-3059, 2845-3058, 2851-3061, 2851-3075, 2852-3027, 2852-3046, 2861-3075, 2865-3059, 2865-3075, 2867-3037, 2868-3075, 2876-3046, 2876-3056, 2876-3059, 2876-3073, 2877-3056, 2878-3039, 2879-3058, 2881-2987, 2883-3063, 2885-3032, 2885-3075, 2887-3056, 2887-3058, 2887-3059, 2887-3073, 2887-3075,
	2890-2989, 2892-3058, 2894-3075, 2895-3042, 2908-3060, 2949-3075, 2998-3057, 2998-3060, 2998-3063, 2998-3064, 2999-3075
108/2321054CB1/ 849	1-96, 1-109, 1-134, 1-273, 100-364, 100-637, 100-656, 100-691, 100-700, 100-831, 100-834, 100-848, 101-453, 132-842, 174-823, 200-731, 200-784, 213-780, 246-833, 273-573, 278-849, 287-796, 297-831, 323-537, 323-553, 334-838, 355-642, 360-587, 361-696, 382-652, 554-776

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
109/2796034CB1/ 2659	1-790, 49-773, 85-773, 118-773, 300-773, 326-773, 498-773, 544-799, 563-804, 608-810, 709-1077, 801-1077, 947-1142, 958-1273, 958-1538, 991-1142, 1008-1219, 1262-2658, 1263-1551, 1408-1997, 1479-1860, 1715-2350, 2017-2618, 2156-2306, 2187-2659
110/4413112CB1/ 2205	1-268, 1-289, 1-622, 10-245, 10-254, 10-358, 10-433, 10-438, 15-263, 367-623, 367-779, 367-816, 367-884, 367-889, 367-897, 367-942, 367-944, 367-947, 367-955, 367-1001, 367-1010, 367-1026, 367-1064, 367-1089, 370-982, 372-464, 382-1075, 458-737, 458-1128, 653-1316, 713-981, 769-1212, 802-1467, 823-1343, 901-1468, 949-1164, 949-1167, 950-1323, 964-1578, 1031-1397, 1034-1650, 1040-1500, 1094-1394, 1129-1804, 1310-1503, 1362-1621, 1416-1512, 1436-1691, 1462-1667, 1462-1717, 1462-1964, 1467-1778, 1478-1619, 1504-1800, 1545-1912, 1611-1837, 1611-2115, 1657-1911, 1703-2205, 1762-2018
111/7654832CB1/ 3042	1-391, 1-567, 29-488, 29-678, 97-699, 104-846, 115-227, 153-627, 153-680, 153-684, 153-726, 153-769, 153-776, 648-942, 649-814, 649-1187, 713-1409, 733-1267, 951-1546, 1107-1800, 1207-1869, 1207-1901, 1207-1938, 1207-1945, 1207-1947, 1207-1981, 1207-1982, 1207-1987, 1207-1994, 1207-1996, 1207-1999, 1208-1938, 1208-1943, 1208-2086, 1274-2012, 1600-1872, 1668-1940, 1668-2141, 1668-2353, 1669-2432, 1819-2113, 1858-2031, 2050-2567, 2051-2327, 2120-2973, 2121-2428, 2135-2973, 2139-2975, 2311-2951, 2393-3009, 2412-3042, 2416-3042, 2423-2949, 2439-3039, 2446-2973, 2460-3009, 2483-2728, 2552-2800, 2559-3042, 2630-2862, 2770-3027
112/7503849CB1/ 2112	1-214, 1-216, 1-249, 1-273, 1-293, 2-239, 4-783, 6-301, 6-343, 8-2112, 10-556, 13-311, 14-303, 14-495, 16-264, 16-633, 17-295, 18-248, 18-269, 18-605, 18-707, 19-252, 19-258, 19-287, 19-536, 19-712, 20-272, 20-374, 21-268, 21-292, 21-328, 21-337, 21-343, 21-444, 22-325, 22-416, 22-703, 23-298, 23-540, 23-641, 24-219, 24-456, 25-687, 25-709, 27-335, 28-304, 29-597, 35-626, 37-665, 37-674, 38-538, 38-704, 39-570, 39-625, 40-631, 40-645, 41-685, 41-784, 42-657, 42-757, 43-301, 45-592, 48-630, 48-708, 51-344, 51-418, 51-421, 53-337, 57-489, 58-386, 59-340, 60-311, 61-418, 62-307, 67-387, 89-333, 115-735, 139-404, 141-338, 141-410, 145-383, 164-440, 165-386, 167-414, 180-464, 186-546, 186-607, 211-466, 218-413, 218-721, 219-539, 220-688, 221-310, 221-446, 230-520, 230-556, 232-485, 235-531, 236-770, 239-874, 278-497, 288-765, 297-748, 314-737, 329-508, 334-580, 384-510, 531-1073, 563-833, 656-945, 674-907, 698-1310, 882-1475, 948-1177, 950-1399, 969-1314, 980-1554, 1012-1252, 1016-1515, 1028-1678, 1030-1298, 1035-1615, 1045-1301, 1047-1501, 1048-1284, 1048-1344, 1056-1684, 1058-1328, 1058-1330, 1058-1336, 1060-1298, 1062-1799, 1066-1582, 1080-1643, 1083-1800, 1091-1668, 1091-1712,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
112 (cont.)	1092-1335, 1098-1657, 1101-1826, 1109-1560, 1109-1624, 1126-1590, 1134-1753, 1137-1590, 1148-1587, 1152-1586, 1153-1715, 1159-1762, 1168-1314, 1170-1437, 1170-1735, 1172-1789, 1172-1871, 1177-1431, 1177-1474, 1177-1791, 1179-1651, 1190-1636, 1192-1713, 1204-1466, 1205-1628, 1206-1470, 1206-1586, 1219-1584, 1221-1489, 1225-1628, 1226-1468, 1232-1590, 1234-1467, 1234-1632, 1240-1590, 1240-1800, 1241-1497, 1241-1537, 1241-1587, 1242-1712, 1244-1485, 1246-1427, 1246-1491, 1248-1836, 1249-1708, 1249-1873, 1253-1712, 1255-1865, 1260-1546, 1266-1712, 1268-1592, 1268-1873, 1270-1497, 1270-1712, 1275-1717, 1275-1718, 1277-1647, 1284-1541, 1295-1826, 1298-1873, 1299-1873, 1300-1715, 1300-1871, 1300-1873, 1302-1873, 1303-1710, 1303-1785, 1307-1871, 1308-1870, 1316-1536, 1316-1868, 1319-1718, 1321-1858, 1324-1590, 1324-1864, 1325-1866, 1343-1872, 1344-1463, 1345-1583, 1360-1586, 1360-1715, 1379-1676, 1382-1873, 1402-1873, 1415-1873, 1424-1494, 1424-1873, 1428-1590, 1428-1729, 1428-1873, 1429-1873, 1441-1712, 1441-1872, 1453-1673, 1456-1873, 1458-1873, 1461-2112, 1462-1873, 1463-1872, 1464-1850, 1464-1873, 1476-1871, 1476-1872, 1479-1736, 1483-
	1834, 1490-1570, 1490-1873, 1493-1811, 1496-1872, 1502-1861, 1520-1796, 1521-1820, 1528-1658, 1530-1873, 1533-1873, 1534-1822, 1534-1830, 1534-1872, 1548-1872, 1554-1873, 1556-1813, 1557-1872, 1563-1704, 1564-1872, 1571-1872, 1573-1873, 1599-1868, 1601-1844, 1601-1845, 1624-1873, 1701-1873, 1748-1859, 1814-1864

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
57	2867236CB1	KIDNNOT20
58	1294096CB1	PROSTUS23
59	7238537CB1	BRAITUT03
60	7494391CB1	SINTNOT18
61	6451054CB1	CORPNOT02
62	7494592CB1	UTRSTMR01
63	5202657CB1	HEARFET01
64	2013529CB1	TESTNOT03
65	3841351CB1	LIVRNOT21
66	152116CB1	BRAITDR02
67	2381031CB1	PROSBPS05
68	2511371CB1	CONUTUT01
69	8068623CB1	TONSDIC01
70	677977CB1	BRABDIE02
71	1661472CB1	MLP000032
72	1748508CB1	PROSNON01
73	2159545CB1	PLACFEB01
74	8560269CB1	NEUTFMT01
75	8710302CB1	THYMNOE02
76	6778214CB1	THYRTUT03
77	258383CB1	BRSTNOT17
78	2804937CB1	STOMTDE01
80	2073751CB1	TLYMUNT01
81	3178841CB1	UTRSTUE01
82	3674807CB1	PLACNOT07
83	1794922CB1	PROSTUT05
84	1795509CB1	PROSTUT05
85	2017180CB1	BLADTUT08
86	219442CB1	THYMNOT05
87	2597459CB1	BRACNOK02
88	2783863CB1	BRAHTDR03
89	2902971CB1	BRAYDIN03
90	368660CB1	BRSTNOT01
91	2804990CB1	BRAITUT02
92	168571CB1	BRAUNOR01
93	1286391CB1	MENITUT03
94	2007684CB1	LUNGNON03
95	2227040CB1	BRSTNOT03
96	4346130CB1	BRAUNOR01
98	7472392CB1	BRALNON02
99	4028960CB1	KIDNFEC01
100	8227004CB1	PROSNOT19
101	3044763CB1	HEAANOT01
102	4044519CB1	LUNGNOT35
103	71351918CB1	BRSTTUT01
104	8109363CB1	BRAINOT19
105	1272746CB1	TESTTUT02
106	1839974CB1	OVARDIR01
107	1877336CB1	ADRETUR01

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
108	2321054CB1	BRSTNOT13
109	2796034CB1	LUNGDIN02
110	4413112CB1	MONOTXT01
111	7654832CB1	BLADTUT06
112	7503849CB1	KIDNNOT09

Table 6

Library	Vector	Library Description
ADRETUR01	PCDNA2.1	This random primed library was constructed using RNA isolated from left upper pole, adrenal gland tumor tissue removed from a 52-year-old Caucasian male during nephroureterectomy and local destruction of renal lesion. Pathology indicated grade 3 adrenal cortical carcinoma forming a mass that infiltrated almost the whole adrenal parenchyma and extended to adjacent adipose tissue. A metastatic tumor nodule was identified in the hilar region. The renal vein was infiltrated by tumor and the neoplastic process was present at the resection margin of the renal vein. Fragments of adrenal cortical carcinoma and thrombus were found in the inferior vena cava. Patient history included abnormal weight loss. Family history included skin cancer, type I diabetes, and neurotic depression.
BLADTUT06	pINCY	Library was constructed using RNA isolated from bladder tumor tissue removed from the posterior bladder wall of a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, and gastrostomy. Pathology indicated grade 3 transitional cell carcinoma in the left lateral bladder wall. The remaining bladder showed marked cystitis with scattered microscopic foci of transitional cell carcinoma in situ. Patient history included angina, emphysema and tobacco use. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
BLADTUT08	pINCY	Library was constructed using RNA isolated from bladder tumor tissue removed from a 72-year-old Caucasian male during a radical cystectomy and prostatectomy. Pathology indicated an invasive grade 3 (of 3) transitional cell carcinoma in the right bladder base. Patient history included pure hypercholesterolemia and tobacco abuse. Family history included myocardial infarction, cerebrovascular disease, and brain cancer.
BRABDIE02	pINCY	This 5' biased random primed library was constructed using RNA isolated from diseased cerebellum tissue removed from the brain of a 57-year-old Caucasian male who died from a cerebrovascular accident. Serologies were negative. Patient history included Huntington's disease, emphysema, and tobacco abuse (3-4 packs per day, for 40 years).

Table 6

Library	Vector	Library Description
BRACNOK02	PSPORT1	This amplified and normalized library was constructed using RNA isolated from posterior cingulate tissue removed from an 85-year-old Caucasian female who died from myocardial infarction and retroperitoneal hemorrhage. Pathology indicated atherosclerosis, moderate to severe, involving the circle of Willis, middle cerebral, basilar and vertebral arteries; infarction, remote, left dentate nucleus; and amyloid plaque deposition consistent with age. There was mild to moderate leptomeningeal fibrosis, especially over the convexity of the frontal lobe. There was mild generalized atrophy involving all lobes. The white matter was mildly thinned. Cortical thickness in the temporal lobes, both maximal and minimal, was slightly reduced. The substantia nigra pars compacta appeared mildly depigmented. Patient history included COPD, hypertension, and recurrent deep venous thrombosis. 6.4 million independent clones from this amplified library were normalized in one round using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791.
BRAHTDR03	PCDNA2.1	This random primed library was constructed using RNA isolated from archaecortex, anterior hippocampus tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.
BRAINOT19	pINCY	Library was constructed using RNA isolated from diseased brain tissue removed from the left frontal lobe of a 27-year-old Caucasian male during a brain lobectomy. Pathology indicated a focal deep white matter lesion, characterized by marked gliosis, calcifications, and hemosiderin-laden macrophages, consistent with a remote perinatal injury. This tissue also showed mild to moderate generalized gliosis, predominantly subpial and subcortical, consistent with chronic seizure disorder. The left temporal lobe, including the mesial temporal structures, showed focal, marked pyramidal cell loss and gliosis in hippocampal sector CA1, consistent with mesial temporal sclerosis. GFAP was positive for astrocytes. The patient presented with intractable epilepsy, focal epilepsy, hemiplegia, and an unspecified brain injury. Patient history included cerebral palsy, abnormality of gait, and depressive disorder. Family history included brain cancer.

Table 6

Library	Vector	Library Description
BRAITDR02	PCDNA2.1	This random primed library was constructed using RNA isolated from allocortex, neocortex, anterior and frontal cingulate tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.
BRAITUT02	PSPORT1	Library was constructed using RNA isolated from brain tumor tissue removed from the frontal lobe of a 58-year-old Caucasian male during excision of a cerebral meningeal lesion. Pathology indicated a grade 2 metastatic hypernephroma. Patient history included a grade 2 renal cell carcinoma, insomnia, and chronic airway obstruction. Family history included a malignant neoplasm of the kidney.
BRAITUT03	PSPORT1	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 17-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a grade 4 fibrillary giant and small-cell astrocytoma. Family history included benign hypertension and cerebrovascular disease.
BRALNON02	pINCY	This thalamus tissue library was constructed from 4.24 million independent clones from a thalamus tissue library. Starting RNA was made from thalamus tissue removed from a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. Microscopically, the cerebral hemisphere revealed moderate fibrosis of the leptomeninges with focal calcifications. There was evidence of shrunken and slightly eosinophilic pyramidal neurons throughout the cerebral hemispheres. Scattered throughout the cerebral cortex, there were multiple small microscopic areas of cavitation with surrounding gliosis. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly and an enlarged spleen and liver. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.

Table 6

Library	Vector	Library Description
BRAUNOR01	pINCY	This random primed library was constructed using RNA isolated from striatum, globus pallidus and posterior putamen tissue removed from an 81-year-old Caucasian female who died from a hemorrhage and ruptured thoracic aorta due to atherosclerosis. Pathology indicated moderate atherosclerosis involving the internal carotids, bilaterally; microscopic infarcts of the frontal cortex and hippocampus; and scattered diffuse amyloid plaques and neurofibrillary tangles, consistent with age. Grossly, the leptomeninges showed only mild thickening and hyalinization along the superior sagittal sinus. The remainder of the leptomeninges was thin and contained some congested blood vessels. Mild atrophy was found mostly in the frontal poles and lobes, and temporal lobes, bilaterally. Microscopically, there were pairs of Alzheimer type II astrocytes within the deep layers of the neocortex. There was increased satellitosis around neurons in the deep gray matter in the middle frontal cortex. The amygdala contained rare diffuse plaques and neurofibrillary tangles. The posterior hippocampus contained a microscopic area of cystic cavitation with hemosiderin-laden macrophages surrounded by
		reactive gliosis. Patient history included sepsis, cholangitis, post-operative atelectasis, pneumonia CAD, cardiomegaly due to left ventricular hypertrophy, splenomegaly, arteriolonephrosclerosis, nodular colloidal goiter, emphysema, CHF, hypothyroidism, and peripheral vascular disease.
BRAYDIN03	pINCY	This normalized library was constructed from 6.7 million independent clones from a brain tissue library. Starting RNA was made from RNA isolated from diseased hypothalamus tissue removed from a 57-year-old Caucasian male who died from a cerebrovascular accident. Patient history included Huntington's disease and emphysema. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 -hours/round) reannealing hybridization was used. The library was linearized and recircularized to select for insert containing clones.
BRSTNOT01	PBLUESCRIPT	Library was constructed using RNA isolated from the breast tissue of a 56-year-old Caucasian female who died in a motor vehicle accident.
BRSTNOT03	PSPORT1	Library was constructed using RNA isolated from diseased breast tissue removed from a 54-year-old Caucasian female during a bilateral radical mastectomy. Pathology for the associated tumor tissue indicated residual invasive grade 3 mammary ductal adenocarcinoma. Patient history included kidney infection and condyloma acuminatum. Family history included benign hypertension, hyperlipidemia and a malignant neoplasm of the colon.

Table 6

Library	Vector	Library Description
BRSTNOT13	pINCY	Library was constructed using RNA isolated from breast tissue removed from a 36-year-old Caucasian female during bilateral simple mastectomy. Patient history included a breast neoplasm, depressive disorder, hyperlipidemia, and a chronic stomach ulcer. Family history included a cardiovascular and cerebrovascular disease; hyperlipidemia; skin, breast, esophageal, bladder, and bone cancer; and Hodgkin's lymphoma.
BRSTNOT17	pINCY	Library was constructed using RNA isolated from breast tissue removed from a 46-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated invasive grade 3, nuclear grade 2 adenocarcinoma, ductal type. An intraductal carcinoma component, non-comedo, comprised approximately 50% of the neoplasm, including the lactiferous ducts. Angiolymphatic involvement was present, and metastatic adenocarcinoma was present in 7 of 10 axillary lymph nodes. The largest nodal metastasis measured 3 cm, and focal extracapsular extension was identified. Family history included atherosclerotic coronary artery disease, type II diabetes, cerebrovascular disease, and depressive disorder.
BRSTTUT01	PSPORT1	Library was constructed using RNA isolated from breast tumor tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated invasive grade 4 mammary adenocarcinoma of mixed lobular and ductal type, extensively involving the left breast. The tumor was identified in the deep dermis near the lactiferous ducts with extracapsular extension. Seven mid and low and five high axillary lymph nodes were positive for tumor. Proliferative fibrocysytic changes were characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Patient history included atrial tachycardia, blood in the stool, and a benign breast neoplasm. Family history included benign hypertension, atherosclerotic coronary artery disease, cerebrovascular disease, and depressive disorder.
CONUTUT01	pINCY	Library was constructed using RNA isolated from sigmoid mesentery tumor tissue obtained from a 61-year-old female during a total abdominal hysterectomy and bilateral salpingo-oophorectomy with regional lymph node excision. Pathology indicated a metastatic grade 4 malignant mixed mullerian tumor present in the sigmoid mesentery at two sites.
CORPNOT02	pINCY	Library was constructed using RNA isolated from diseased corpus callosum tissue removed from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease.

Table 6

Library	Vector	Library Description
HEAANOT01	pINCY	Library was constructed using RNA isolated from right coronary and right circumflex coronary artery tissue removed from the explanted heart of a 46-year-old Caucasian male during a heart transplantation. Patient history included myocardial infarction from total occlusion of the left anterior descending coronary artery, atherosclerotic coronary artery disease, hyperlipidemia, myocardial ischemia, dilated cardiomyopathy, left ventricular dysfunction, and tobacco abuse. Previous surgeries included cardiac catheterization. Family history included atherosclerotic coronary artery disease.
HEARFET01	pINCY	Library was constructed using RNA isolated from heart tissue removed from a Hispanic male fetus, who died at 18 weeks' gestation.
KIDNFEC01	PBLUESCRIPT	Library was constructed using RNA isolated from kidney tissue removed from a pool of twelve Caucasian male and female fetuses that were spontaneously aborted at 19-23 weeks' gestation.
KIDNNOT09	pINCY	Library was constructed using RNA isolated from the kidney tissue of a Caucasian male fetus, who died at 23 weeks' gestation.
KIDNNOT20	pINCY	Library was constructed using RNA isolated from left kidney tissue removed from a 43-year-old Caucasian male during nephroureterectomy, regional lymph node excision, and unilateral left adrenalectomy. Pathology for the associated tumor tissue indicated a grade 2 renal cell carcinoma. Family history included atherosclerotic coronary artery disease.
LIVRNOT21	pINCY	Library was constructed using RNA isolated from liver tissue removed from a 29-year-old Caucasian male who died from massive head injury due to a motor vehicle accident. Serology was positive for cytomegalovirus.
LUNGDIN02	pINCY	This normalized lung tissue library was constructed from 7.6 million independent clones from a diseased lung tissue library. Starting RNA was made from RNA isolated from diseased lung tissue. Pathology indicated ideopathic pulmonary disease. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
LUNGNON03	PSPORT1	This normalized library was constructed from 2.56 million independent clones from a lung tissue library. RNA was made from lung tissue removed from the left lobe of a 58-year-old Caucasian male during a segmental lung resection. Pathology for the associated tumor tissue indicated a metastatic grade 3 (of 4) osteosarcoma. Patient history included soft tissue cancer, secondary cancer of the lung, prostate cancer, and an acute duodenal ulcer with hemorrhage. Patient also received radiation therapy to the retroperitoneum. Family history included prostate cancer, breast cancer, and acute leukemia. The normalization and hybridization conditions were adapted from Soares et al., PNAS (1994) 91:9228; Swaroop et al., NAR (1991) 19:1954; and Bonaldo et al., Genome Research (1996) 6:791.

Table 6

Library	Vector	Library Description
LUNGNOT35	pINCY	Library was constructed using RNA isolated from lung tissue removed from a 62-year-old Caucasian female. Pathology for the associated tumor tissue indicated a grade 1 spindle cell carcinoid forming a nodule. Patient history included depression, thrombophlebitis, and hyperlipidemia. Family history included cerebrovascular disease, atherosclerotic coronary artery disease, breast cancer, colon cancer,type II diabetes, and malignant skin melanoma.
MENITUT03	pINCY	Library was constructed using RNA isolated from brain meningioma tissue removed from a 35-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a benign neoplasm in the right cerebellopontine angle of the brain. Patient history included hypothyroidism. Family history included myocardial infarction and breast cancer .
MLP000032	PCR2-TOPOTA	Library was constructed using pooled cDNA from different donors. cDNA was generated using mRNA isolated from the following: aorta, cerebellum, lymph nodes, muscle, tonsil (lymphoid hyperplasia), bladder tumor (invasive grade 3 transitional cell carcinoma.), breast (proliferative fibrocystic changes without atypia characterized by epithelial ductal hyperplasia, testicle tumor (embryonal carcinoma), spleen, ovary, parathyroid, ileum, breast skin, sigmoid colon, penis tumor (fungating invasive grade 4 squamous cell carcinoma), fetal lung,, breast, fetal small intestine, fetal liver, fetal pancreas, fetal lung, fetal skin, fetal penis, fetal bone, fetal ribs, frontal brain tumor (grade 4 gemistocytic astrocytoma), ovary (stromal hyperthecosis), bladder, bladder tumor (invasive grade 3 transitional cell carcinoma), stomach, lymph node tumor (metastatic basaloid squamous cell carcinoma), tonsil (reactive lymphoid hyperplasia), periosteum from the tibia, fetal brain, fetal spleen, uterus tumor, endometrial (grade 3 adenosquamous carcinoma), seminal vesicle, liver, aorta, adrenal gland, lymph node (metastatic grade 3 squamous cell carcinoma), glossal muscle, esophagus,
		esophagus tumor (invasive grade 3 adenocarcinoma), ileum, pancreas, soft tissue tumor from the skull (grade 3 ependymoma), transverse colon, (benign familial polyposis), rectum tumor (grade 3 colonic adenocarcinoma), rib tumor, (metastatic grade 3 osteosarcoma), lung, heart, placenta, thymus, stomach, spleen (splenomegaly with congestion), uterus, cervix (mild chronic cervicitis with focal squamous metaplasia), spleen tumor (malignant lymphoma, diffuse large cell type B-cell phenotype with abundant reactive T-cells and marked granulomatous response), umbilical cord blood mononuclear cells, upper lobe lung tumor, (grade 3 squamous cell carcinoma), endometrium (secretory phase), liver, liver tumor (metastatic grade 2 neuroendocrine carcinoma), colon, umbilical cord blood, Th1 cells, nonactivated, umbilical cord blood, Th2 cells, nonactivated, coronary artery endothelial cells (untreated), coronary artery smooth muscle cells, (untreated), coronary artery smooth muscle cells (treated with TNF & IL-1 10ng/ml each for 20 hours), bladder (mild chronic cystitis), epiglottis, breast skin, small intestine, fetal prostate stroma fibroblasts, prostate epithelial cells (PrEC cells),

Table 6

Library	Vector	Library Description
MLP000032 (cont.)		fetal adrenal glands, fetal liver, kidney transformed embryonal cell line (293-EBNA) (untreated), kidney transformed embryonal cell line (293-EBNA) (treated with 5Aza-2deoxycytidine for 72 hours), mammary epithelial cells, (HMEC cells), peripheral blood monocytes (treated with IL-10 at time 0, 10ng/ml, LPS was added at 1 hour at 5ng/ml. Incubation 24 hours), peripheral blood monocytes (treated with anti-IL-10 at time 0, 10ng/ml, LPS was added at 1 hour at 5ng/ml. Incubation 24 hours), spinal cord, base of medulla (Huntington's chorea), thigh and arm muscle (ALS), breast skin fibroblast (untreated), breast skin fibroblast (treated with 9CIS Retinoic Acid 1μM for 20 hours), breast skin fibroblast (treated with TNF-alpha & IL-1 beta, 10ng/ml each for 20 hours), fetal liver mast cells, hematopoietic (Mast cells prepared from human fetal liver hematopoietic progenitor cells (CD34+ stem cells) cultured in the presence of hIL-6 and hSCF for 18 days), epithelial layer of colon, bronchial epithelial cells (treated for 20hours with 20% smoke conditioned media), lymph node, pooled peripheral blood mononuclear cells (untreated), pooled brain segments: striatum, globus
		pallidus and posterior putamen (Alzheimer's Disease), pituitary gland, umbilical cord blood, CD34+ derived dendritic cells (treated with SCF, GM-CSF & TNF alpha, 13 days), umbilical cord blood, CD34+ derived dendritic cells (treated with SCF, GM-CSF & TNF alpha, 13 days followed by PMA/Ionomycin for 5 hours), small intestine, rectum, bone marrow neuroblastoma cell line (SH-SY5Y cells, treated with 6-Hydroxydopamine 100 uM for 8 hours), bone marrow, neuroblastoma cell line (SH-SY5Y cells, untreated), brain segments from one donor: amygdala, entorhinal cortex, globus pallidus, substantia innominata, striatum, dorsal caudate nucleus, dorsal putamen, ventral nucleus accumbens, archaecortex (hippocampus anterior and posterior), thalamus, nucleus raphe magnus, periaqueductal gray, midbrain, substantia nigra, and dentate nucleus, pineal gland (Alzheimer's Disease), preadipocytes (untreated), preadipocytes (treated with a peroxisome proliferator-activated receptor gamma agonist, 1microM, 4 hours), pooled prostate (adenofibromatous hyperplasia), pooled kidney, pooled adipocytes (untreated), pooled adipocytes (treated with human insulin),

Table 6

Library	Vector	Library Description
MLP000032 (cont.)		pooled mesentaric and abdomenal fat, pooled adrenal glands, pooled thyroid (normal and adenomatous hyperplasia), pooled spleen (normal and with changes consistent with idiopathic thrombocytopenic purpura), pooled right and left breast, pooled lung, pooled nasal polyps, pooled fat, pooled synovium (normal and rhumatoid arthritis), pooled brain (meningioma, gemistocytic astrocytoma. and Alzheimer's disease), pooled fetal colon, pooled colon: ascending, descending (chronic ulcerative colitis), and rectal tumor (adenocarcinoma), pooled esophagus, normal and tumor (invasive grade 3 adenocarcinoma), pooled breast skin fibroblast (one treated w/9CIS Retinoic Acid and the other with TNF-alpha & IL-1 beta), pooled gallbladder (acute necrotizing cholecystitis with cholelithiasis (clinically hydrops), acute hemorrhagic cholecystitis with cholelithiasis, chronic cholecystitis and cholelithiasis), pooled fetal heart, (Patau's and fetal demise), pooled neurogenic tumor cell line, SK-N-MC, (neuroepitelioma, metastasis to supra-orbital area, untreated) and neuron, NT-2 cell line, (treated with mouse leptin at 1 μg/ml and 9cis retinoic acid at 3.3 μM
		for 6 days), pooled ovary (normal and polycystic ovarian disease), pooled prostate, (adenofibromatous hyperplasia), pooled seminal vesicle, pooled small intestine, pooled fetal small intestine, pooled stomach and fetal stomach, prostate epithelial cells, pooled testis (normal and embryonal carcinoma), pooled uterus, pooled uterus tumor (grade 3 adenosquamous carcinoma and leiomyoma), pooled uterus, endometrium, and myometrium, (normal and adenomatous hyperplasia with squamous metaplasia and focal atypia), pooled brain: (temporal lobe meningioma, cerebellum and hippocampus (Alzheimer's Disease), pooled skin, fetal lung, adrenal tumor (adrenal cortical carcinoma), prostate tumor (adenocarcinoma), fetal heart, fetal small intestine, ovary tumor (mucinous cystadenoma), ovary, ovary tumor (transitional cell carcinoma), disease prostate (adenofibromatous hyperplasia), fetal colon, uterus tumor (leiomyoma), temporal brain, submandibular gland, colon tumor (adenocarcinoma), ascending and transverse colon, ovary tumor (endometrioid carcinoma), lung tumor (squamous cell carcinoma), fetal brain, fetal lung, ureter tumor (transitional cell carcinoma),

Table 6

Library	Vector	Library Description
MLP000032 (cont.)		untreated HNT cells, para-aortic soft tissue, testis, seminal vesicle, diseased ovary (endometriosis), temporal lobe, myometrium, diseased gallbladder (cholecystitis, cholelithiasis), placenta, breast tumor (ductal adenocarcinoma), breast, lung tumor (liposarcoma), endometrium, abdominal fat, cervical spine dorsal root ganglion, thoracic spine dorsal root ganglion, diseased thyroid (adenomatous hyperplasia), liver, kidney, fetal liver, NT-2 cells (treated with mouse leptin and 9cis RA), K562 cells (treated with 9cis RA), cerebellum, corpus callosum, hypothalamus, fetal brain astrocytes (treated with TNFa and IL-1b), inferior parietal cortex, posterior hippocampus, pons, thalamus, C3A cells (untreated), C3A cells (treated with 3-methylcholanthrene), testis, colon epithelial layer, pooled prostate, pooled liver, substantia nigra, thigh muscle, rib bone, fallopian tube tumor (endometrioid and serous adenocarcinoma), diseased lung (idiopathic pulmonary disease), cingulate anterior allocortex and neocortex, cingulate posterior allocortex, auditory neocortex, frontal neocortex, orbital inferior neocortex, parietal superior neocortex, visual primary neocortex, dentate nucleus, posterior cingulate,
		cerebellum, vermis, inferior temporal cortex, medulla, posterior parietal cortex, colon polyp, pooled breast, anterior and posterior hippocampus, mesenteric and abdominal fat, pooled esophagus, pooled fetal kidney, pooled fetal liver, ileum, small intestine, pooled gallbladder, frontal and superior temporal cortex, pooled ovary, pooled endometrium, pooled prostate, pooled kidney, fetal femur, sacrum tumor (giant cell tumor), pooled kidney and kidney tumor (renal cell carcinoma clear-cell type), pooled liver and liver tumor (neuroendocrine carcinoma), pooled fetal liver, pooled lung, fetal pancreas, pancreas, parotid gland, parotid tumor (sebaceous lymphadenoma), retroperitoneal and suprglottic soft tissue, spleen, fetal spleen, spleen tumor (malignant lymphoma), diseased spleen (idiopathic thrombocytopenic purpura), parathyroid, thyroid, thymus, tonsil ureter tumor (transitional cell carcinoma), pooled adrenal gland and adrenal tumor (pheochromocytoma), pooled lymph node tumor (Hodgkin's disease and metastatic adenocarcinoma), pooled neck and calf muscles, and pooled bladder
MONOTXT01	pINCY	Library was constructed using RNA isolated from treated monocytes from peripheral blood obtained from a 42-year-old female. The cells were treated with anti IL-10 and LPS.
NEUTFMT01	PBLUESCRIPT	Library was constructed using total RNA isolated from peripheral blood granulocytes collected by density gradient centrifugation through Ficoll-Hypaque. The cells were isolated from buffy coat units obtained from unrelated male and female donors. Cells were cultured in 10 nM fMLP for 30 minutes, lysed in GuSCN, and spun through CsCl to obtain RNA for library construction. Because this library was made from total RNA, it has an unusually high proportion of unique singleton sequences, which may not all come from polyA RNA species.

Table 6

Library	Vector	Library Description
OVARDIR01	PCDNA2.1	This random primed library was constructed using RNA isolated from right ovary tissue removed from a 45-year-old Caucasian female during total abdominal hysterectomy, bilateral salpingo-oophorectomy, vaginal suspension and fixation, and incidental appendectomy. Pathology indicated stromal hyperthecosis of the right and left ovaries. Pathology for the matched tumor tissue indicated a dermoid cyst (benign cystic teratoma) in the left ovary. Multiple (3) intramural leiomyomata were identified. The cervix showed squamous metaplasia. Patient history included metrorrhagia, female stress incontinence, alopecia, depressive disorder, pneumonia, normal delivery, and deficiency anemia. Family history included benign hypertension, atherosclerotic coronary artery disease, hyperlipidemia, and primary tuberculous complex.
PLACFEB01	pINCY	Library was constructed using pooled cDNA from two different donors. cDNA was generated using RNA isolated from placenta tissue removed from a Caucasian fetus (donor A), who died after 16 weeks' gestation from fetal demise and hydrocephalus; and a Caucasian male fetus (donor B), who died after 18 weeks' gestation from fetal demise. Patient history included umbilical cord wrapped around the head (3 times) and the shoulders (1 time) in donor A. Serology was positive for anti-CMV in donor A. Family history included multiple pregnancies and live births, and an abortion in donor A.
PLACNOT07	pINCY	Library was constructed using RNA isolated from placental tissue removed from a Caucasian fetus, who died after 16 weeks' gestation from fetal demise and hydrocephalus. Serology was positive for anti-CMV (cytomegalovirus).
PROSBPS05	pINCY	This subtracted prostate tissue library was constructed using 4.48x10e5 clones from diseased prostate tissue and was subjected to two rounds of subtraction hybridization with 1.56 million clones from a breast tissue library. The starting library for subtraction was constructed using RNA isolated from diseased prostate tissue removed from a 70-year-old Caucasian male during a radical prostatectomy and closed prostatic biopsy. Pathology indicated benign prostatic hypertrophy. Pathology for the matched tumor tissue indicated adenocarcinoma. The patient presented with elevated prostate specific antigen and induration. Patient history included benign hypertension, gastrointestinal bleed, cardiac dysrhythmia, cardiac arrest, hyperlipidemia, alcohol abuse and fractured mandible. Previous surgeries included splenectomy, cholecystectomy and inguinal hernia repair. Patient medications included Verapamil and antacids. Family history included benign hypertension, myocardial infarction and coronary atherosclerosis in the mother; tobacco abuse and lung cancer in the father; tobacco abuse, cerebrovascular accident and lung cancer in the sibling(s). The
		hybridization probe for subtraction was derived from a similarly constructed library from RNA isolated from nontumorous breast tissue from a different donor. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR 19 (1991):1954 and Bonaldo, et al. Genome Research 6 (1996): 791.

Table 6

Library	Vector	Library Description
PROSNON01	PSPORT1	This normalized prostate library was constructed from 4.4 M independent clones from a prostate library. Starting RNA was made from prostate tissue removed from a 28-year-old Caucasian male who died from a self-inflicted gunshot wound. The normalization and hybridization conditions were adapted from Soares, M.B. et al. (1994) Proc. Natl. Acad. Sci. USA 91:9228-9232, using a longer (19 hour) reannealing hybridization period.
PROSNOT19	pINCY	Library was constructed using RNA isolated from diseased prostate tissue removed from a 59-year-old Caucasian male during a radical prostatectomy with regional lymph node excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 3+3). The patient presented with elevated prostate-specific antigen (PSA). Patient history included colon diverticuli, asbestosis, and thrombophlebitis. Previous surgeries included a partial colectomy. Family history included benign hypertension, multiple myeloma, hyperlipidemia and rheumatoid arthritis.
PROSTUS23	pINCY	This subtracted prostate tumor library was constructed using 10 million clones from a pooled prostate tumor library that was subjected to 2 rounds of subtractive hybridization with 10 million clones from a pooled prostate tissue library. The starting library for subtraction was constructed by pooling equal numbers of clones from 4 prostate tumor libraries using mRNA isolated from prostate tumor removed from Caucasian males at ages 58 (A), 61 (B), 66 (C), and 68 (D) during prostatectomy with lymph node excision. Pathology indicated adenocarcinoma in all donors. History included elevated PSA, induration and tobacco abuse in donor A; elevated PSA, induration, prostate hyperplasia, renal failure, osteoarthritis, renal artery stenosis, benign HTN, thrombocytopenia, hyperlipidemia, tobacco/alcohol abuse and hepatitis C (carrier) in donor B; elevated PSA, induration, and tobacco abuse in donor C; and elevated PSA, induration, hypercholesterolemia, and kidney calculus in donor D. The hybridization probe for subtraction was constructed by pooling equal numbers of cDNA clones from 3 prostate tissue libraries derived from prostate tissue, prostate epithelial cells,
		and fibroblasts from prostate stroma from 3 different donors. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR 19 (1991):1954 and Bonaldo, et al. Genome Research 6 (1996):791.
PROSTUT05	PSPORT1	Library was constructed using RNA isolated from prostate tumor tissue removed from a 69-year-old Caucasian male during a radical prostatectomy. Pathology indicated adenocarcinoma (Gleason grade 3+4). Adenofibromatous hyperplasia was also present. Family history included congestive heart failure, multiple myeloma, hyperlipidemia, and rheumatoid arthritis.
SINTNOT18	pINCY	Library was constructed using RNA isolated from small intestine tissue obtained from a 59-year-old male.

Table 6

Library	Vector	Library Description
STOMTDE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from stomach tissue removed from a 61-year-old Caucasian male during a partial esophagectomy, proximal gastrectomy, pyloromyotomy, and regional lymph node excision. Pathology for the associated tumor indicated an invasive grade 3 adenocarcinoma in the esophagus, extending distally to involve the gastroesophageal junction. The tumor extended through the muscularis to involve periesophageal and perigastric soft tissues. One perigastric and two periesophageal lymph nodes were positive for tumor. There were multiple perigastric and periesophageal tumor implants. The patient presented with deficiency anemia and myelodysplasia. Patient history included hyperlipidemia, and tobacco and alcohol abuse in remission. Previous surgeries included adenotonsillectomy, rhinoplasty, vasectomy, and hemorrhoidectomy. A previous bone marrow aspiration found the marrow to be hypercellular for age and had a cellularity-to-fat ratio of 95:5. The marrow was focally densely fibrotic. Granulocytic precursors were slightly increased with normal maturation. The estimate of blast cells was greater than 5%.
		Megakaryocytes were increased and appeared atypical in clusters. Storage cells and granulomata were absent. Patient medications included Epoetin, Danocrine, Berocca Plus tablets, Selenium, vitamin B6 phosphate, vitamins E & C, and beta carotene. Family history included alcohol abuse, atherosclerotic coronary artery disease, type II diabetes, chronic liver disease, and primary cardiomyopathy in the father; and benign hypertension and cerebrovascular disease in the mother.
TESTNOT03	PBLUESCRIPT	Library was constructed using RNA isolated from testicular tissue removed from a 37-year-old Caucasian male, who died from liver disease. Patient history included cirrhosis, jaundice, and liver failure.
TESTTUT02	pINCY	Library was constructed using RNA isolated from testicular tumor removed from a 31-year-old Caucasian male during unilateral orchiectomy. Pathology indicated embryonal carcinoma.
THYMNOE02	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from thymus tissue removed from a 3-year-old Hispanic male during a thymectomy and closure of a patent ductus arteriosus. The patient presented with severe pulmonary stenosis and cyanosis. Patient history included a cardiac catheterization and echocardiogram. Previous surgeries included Blalock-Taussig shunt and pulmonary valvotomy. The patient was not taking any medications. Family history included benign hypertension, osteoarthritis, depressive disorder, and extrinsic asthma in the grandparent(s).

Table 6

Library	Vector	Library Description
THYMNOT05	pINCY	Library was constructed using RNA isolated from thymus tissue removed from a 3-year-old Hispanic male during a thymectomy and closure of a patent ductus arteriosus. The patient presented with severe pulmonary stenosis and cyanosis. Patient history included a cardiac catheterization and echocardiogram. Previous surgeries included Blalock-Taussig shunt and pulmonary valvotomy. The patient was not taking any medications. Family history included benign hypertension, osteoarthritis, depressive disorder, and extrinsic asthma in the grandparent(s).
THYRTUT03	pINCY	Library was constructed using RNA isolated from benign thyroid tumor tissue removed from a 17-year-old Caucasian male during a thyroidectomy. Pathology indicated encapsulated follicular adenoma forming a circumscribed mass.
TLYMUNT01	pINCY	Library was constructed using RNA isolated from resting allogenic T-lymphocyte tissue removed from an adult (40-50-year old) Caucasian male.
TONSDIC01	PSPORT1	This large size fractionated library was constructed using pooled cDNA from two donors. cDNA was generated using mRNA isolated from diseased left tonsil tissue removed from a 6-year-old Caucasian male (donor A) during adenotonsillectomy and from diseased right tonsil tissue removed from a 9-year-old Caucasian female (donor B) during adenotonsillectomy. Pathology indicated reactive lymphoid hyperplasia, bilaterally (A) and lymphoid hyperplasia (B). The patients presented with sleep apnea (A) and hypertrophy of tonsils, cough, and unspecified nasal and sinus disease (B). Patient history included a bacterial infection (A). Previous surgeries included myringotomy with tube insertion (A). Donor A was not taking any medications and donor B was taking Vancenase. Family history included benign hypertension, myocardial infarction, and atherosclerotic coronary artery disease in the grandparent(s) of donor A; and extrinsic asthma and unspecified allergy in the mother; unspecified allergy in the father; benign hypertension, deficiency anemia, osteoarthritis, extrinsic asthma and unspecified allergy in the grandparent(s) of donor B.
UTRSTMR01	pINCY	Library was constructed using RNA isolated from uterine myometrial tissue removed from a 41-year-old Caucasian female during a vaginal hysterectomy. The endometrium was secretory and contained fragments of endometrial polyps. Pathology for associated tumor tissue indicated uterine leiomyoma. Patient history included ventral hernia and a benign ovarian neoplasm.

Table 6

Library	Vector	Library Description
UTRSTUE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from uterus tumor tissue removed a 37-year-old Black female during myomectomy, dilation and curettage, right fimbrial region biopsy, and incidental appendectomy. Pathology indicated multiple (12) uterine leiomyomata. A fimbrial cyst was identified. The patient presented with deficiency anemia, an umbilical hernia, and premenopausal menorrhagia. Patient history included premenopausal menorrhagia and sarcoidosis of the lung. Previous surgeries included hysteroscopy, dilation and curettage, and an endoscopic lung biopsy. Patient medications included Chromagen and Claritin. Family history included acute myocardial infarction and atherosclerotic coronary artery disease in the father.

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	*ESTs:* Probability value=1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value=1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM, INCY, SMART, and TIGRFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	*PFAM, INCY, SMART, or TIGRFAM hits:* Probability value=1.0E-3 or less *Signal peptide hits:* Score= 0 or greater

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score=120 or greater; Match length=56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
112	7503849	1339126H1	SNP00112738	151	315	A	G	A	T80	n/a	n/a	n/a	n/a
112	7503849	1376303H1	SNP00009519	85	38	C	C	A	noncoding	n/a	n/a	n/a	n/a
112	7503849	1376303H1	SNP00037704	88	35	G	G	A	noncoding	n/a	n/a	n/a	n/a
112	7503849	1422323H1	SNP00061236	167	384	T	C	T	stop103	n/d	n/d	n/d	n/d
112	7503849	1520683H1	SNP00009519	131	39	C	C	A	noncoding	n/a	n/a	n/a	n/a
112	7503849	1520683H1	SNP00037704	134	36	G	G	A	noncoding	n/a	n/a	n/a	n/a
112	7503849	2211818H1	SNP00097061	219	1021	C	C	T	noncoding	n/a	n/a	n/a	n/a
112	7503849	2346103H1	SNP00037704	106	27	G	G	A	noncoding	n/a	n/a	n/a	n/a
112	7503849	2497486H1	SNP00009519	84	41	C	C	A	noncoding	n/a	n/a	n/a	n/a
112	7503849	2497486H1	SNP00037704	87	38	G	G	A	noncoding	n/a	n/a	n/a	n/a
112	7503849	2688833H1	SNP00009519	103	42	C	C	A	noncoding	n/a	n/a	n/a	n/a
112	7503849	2688833H1	SNP00037704	106	39	G	G	A	noncoding	n/a	n/a	n/a	n/a
112	7503849	2691537H1	SNP00009519	228	106	G	G	T	R10	n/a	n/a	n/a	n/a
112	7503849	2691537H1	SNP00037704	225	109	C	C	T	P11	n/a	n/a	n/a	n/a
112	7503849	2984764H1	SNP00037704	84	100	C	C	T	P8	n/a	n/a	n/a	n/a
112	7503849	2998532H1	SNP00092683	42	1103	C	C	T	noncoding	n/a	n/a	n/a	n/a
112	7503849	3083817H1	SNP00092683	97	1102	T	C	T	noncoding	n/a	n/a	n/a	n/a
112	7503849	3083817H1	SNP00097061	15	1020	C	C	T	noncoding	n/a	n/a	n/a	n/a
112	7503849	3163442H1	SNP00097061	252	1015	C	C	T	noncoding	n/a	n/a	n/a	n/a
112	7503849	3184509H1	SNP00009519	79	94	G	G	T	G6	n/a	n/a	n/a	n/a
112	7503849	3184509H1	SNP00037704	82	97	C	C	T	P7	n/a	n/a	n/a	n/a
112	7503849	3184509H1	SNP00112738	288	305	G	G	A	S76	n/a	n/a	n/a	n/a
112	7503849	3319727H1	SNP00092683	71	1101	C	C	T	noncoding	n/a	n/a	n/a	n/a
112	7503849	3360460H1	SNP00009519	85	44	C	C	A	noncoding	n/a	n/a	n/a	n/a
112	7503849	3501445H1	SNP00009519	97	102	T	G	T	C9	n/a	n/a	n/a	n/a
112	7503849	3692522H1	SNP00061236	202	381	C	C	T	P102	n/d	n/d	n/d	n/d
112	7503849	3748594H1	SNP00112738	84	313	A	G	A	H79	n/a	n/a	n/a	n/a
112	7503849	3777748H1	SNP00112738	239	291	G	G	A	G72	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
112	7503849	3877579H1	SNP00009519	47	96	T	G	T	S7	n/a	n/a	n/a	n/a
112	7503849	4205658H1	SNP00061236	49	382	C	C	T	A102	n/d	n/d	n/d	n/d
112	7503849	4541534H1	SNP00112739	225	609	C	C	T	H178	n/d	n/d	n/d	n/d
112	7503849	4575407H1	SNP00092683	54	1100	C	C	T	noncoding	n/a	n/a	n/a	n/a
112	7503849	4640189H1	SNP00112739	206	611	C	C	T	Y178	n/d	n/d	n/d	n/d
112	7503849	4916293H1	SNP00092683	43	1099	C	C	T	noncoding	n/a	n/a	n/a	n/a
112	7503849	4956713H1	SNP00097061	8	1019	C	C	T	noncoding	n/a	n/a	n/a	n/a
112	7503849	5026038H1	SNP00112739	75	602	C	C	T	R175	n/d	n/d	n/d	n/d
112	7503849	5206042H1	SNP00097061	135	851	C	C	T	noncoding	n/a	n/a	n/a	n/a
112	7503849	5872016H1	SNP00061236	159	375	C	C	T	L100	n/d	n/d	n/d	n/d
112	7503849	6732363H1	SNP00112740	410	627	C	C	T	L184	n/a	n/a	n/a	n/a

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-55,

b) a polypeptide consisting essentially of the amino acid sequence of SEQ ID NO:56,

c) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:7-12, SEQ ID NO:14-20, SEQ ID NO:23-24, SEQ ID NO:26-36, SEQ ID NO:38-43, SEQ ID NO:45-46 and SEQ ID NO:48-54,

d) a polypeptide comprising a naturally occurring amino acid sequence at least 93% identical to the amino acid sequence of SEQ ID NO:37,

e) a polypeptide comprising a naturally occurring amino acid sequence at least 94% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:22 and SEQ ID NO:55,

f) a polypeptide comprising a naturally occurring amino acid sequence at least 95% identical to the amino acid sequence of SEQ ID NO:13,

g) a polypeptide comprising a naturally occurring amino acid sequence at least 97% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:6 and SEQ ID NO:44,

h) a polypeptide comprising a naturally occurring amino acid sequence at least 98% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:21 and SEQ ID NO:25,

i) a polypeptide comprising a naturally occurring amino acid sequence at least 99% identical to the amino acid sequence of SEQ ID NO:1,

j) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, and

k) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-56.

2. An isolated polypeptide of claim 1 selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-55 and

b) a polypeptide consisting essentially of the amino acid sequence of SEQ ID NO:56.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method of producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-56.

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

12. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112,

b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:57-112,

c) a polynucleotide complementary to a polynucleotide of a),

d) a polynucleotide complementary to a polynucleotide of b), and

e) an RNA equivalent of a)-d).

13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

18. A composition of claim 17, wherein the polypeptide is selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, and

b) a polypeptide consisting essentially of the amino acid sequence of SEQ ID NO:56.

19. A method for treating a disease or condition associated with decreased expression of functional MDDT, comprising administering to a patient in need of such treatment the composition of claim 17.

20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with decreased expression of functional MDDT, comprising administering to a patient in need of such treatment a composition of claim 21.

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

25. A method for treating a disease or condition associated with overexpression of functional MDDT, comprising administering to a patient in need of such treatment a composition of claim 24.

26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, the method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

b) assessing the activity of the polypeptide of claim 1 in the presence of the test

compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

28. A method of screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,

b) detecting altered expression of the target polynucleotide, and

c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

29. A method of assessing toxicity of a test compound, the method comprising:

a) treating a biological sample containing nucleic acids with the test compound,

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,

c) quantifying the amount of hybridization complex, and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

30. A diagnostic test for a condition or disease associated with the expression of MDDT in a biological sample, the method comprising:

a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and

b) detecting the complex, wherein the presence of the complex correlates with the

presence of the polypeptide in the biological sample.

31. The antibody of claim 11, wherein the antibody is:

a) a chimeric antibody,

b) a single chain antibody,

c) a Fab fragment,

d) a $F(ab')_2$ fragment, or

e) a humanized antibody.

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

33. A method of diagnosing a condition or disease associated with the expression of MDDT in a subject, comprising administering to said subject an effective amount of the composition of claim 32.

34. A composition of claim 32, wherein the antibody is labeled.

35. A method of diagnosing a condition or disease associated with the expression of MDDT in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibodies from the animal, and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56.

37. A polyclonal antibody produced by a method of claim 36.

38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.

39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-56, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
b) isolating antibody producing cells from the animal,
c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
d) culturing the hybridoma cells, and
e) isolating from the culture monoclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56.

40. A monoclonal antibody produced by a method of claim 39.

41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.

42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.

44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56 in a sample, the method comprising:

a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56 in the sample.

45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56 from a sample, the method comprising:

a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and

b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-56.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

a) labeling the polynucleotides of the sample,

b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and

c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.

81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.

82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.

83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.

84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.

85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.

86. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.

87. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.

88. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.

89. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.

90. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.

91. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.

92. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:37.

93. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.

94. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:39.

95. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:40.

96. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:41.

97. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:42.

98. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:43.

99. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:44.

100. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:45.

101. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:46.

102. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:47.

103. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:48.

104. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:49.

105. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:50.

106. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:51.

107. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:52.

108. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:53.

109. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:54.

110. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:55.

111. A polypeptide of claim 1, consisting essentially of the amino acid sequence of SEQ ID NO:56.

112. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:57.

113. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:58.

114. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:59.

115. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:60.

116. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:61.

117. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:62.

118. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:63.

119. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:64.

120. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:65.

121. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:66.

122. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:67.

123. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:68.

124. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:69.

125. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:70.

126. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:71.

127. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:72.

128. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:73.

129. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:74.

130. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:75.

131. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:76.

132. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:77.

133. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:78.

134. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:79.

135. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:80.

136. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:81.

137. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:82.

138. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:83.

139. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:84.

140. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:85.

141. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:86.

142. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:87.

143. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:88.

144. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:89.

145. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:90.

146. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:91.

147. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:92.

148. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:93.

149. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:94.

150. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:95.

151. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:96.

152. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:97.

153. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:98.

154. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:99.

155. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:100.

156. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:101.

157. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:102.

158. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:103.

159. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:104.

160. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:105.

161. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:106.

162. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:107.

163. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:108.

164. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:109.

165. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:110.

166. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:111.

167. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:112.

<110> INCYTE GENOMICS, INC.
THORNTON, Michael
AU-YOUNG, Janice K.
AZIMZAI, Yalda
BANDMAN, Olga
BARROSO, Ines
BAUGHN, Mariah R.
BECHA, Shanya D.
BOROWSKY, Mark L.
DING, Li
DUGGAN, Brendan M.
ELLIOTT, Vicki S.
EMERLING, Brooke M.
FORSYTHE, Ian J.
GANDHI, Ameena R.
GIETZEN, Kimberly J.
GORVAD, Ann E.
GRIFFIN, Jennifer A.
GURURAJAN, Rajagopal
HAFALIZA, April J.A.
RING, Huijun
ISON, Craig H.
JONES, Karen Anne
LAL, Preeti G.
LEE, Ernestine A.
LEE, Sally
LI, Joana X.
LU, Dyung Aina M.
MARQUIS, Joseph
LEHR-MASON, Patricia M.
WALIA, Narinder K.
ARVIZU, Chandra
SANJANWALA, Bharati,
SORNASSE, Thierry
SWARNAKAR, Anita
TANG, Y. Tom
THANGAVELU, Kavitha
TRAN, Bao
TRAN, Uyen K.
WARREN, Bridget A.
XU, Yuming
YAO, Monique G.
YUE, Henry
YUE, Huibin
ZEBARJADIAN, Yeganeh
CHANG, Hsin-Ru

<120> MOLECULES FOR DISEASE DETECTION AND TREATMENT

<130> PF-1063 PCT

<140> To Be Assigned
<141> Herewith

<150> US 60/304,298
<151> 2001-07-09

<150> US 60/305,324
<151> 2001-07-13

<150> US 60/307,003
<151> 2001-07-19

<150> US 60/308,185
<151> 2001-07-27

<150> US 60/310,096
<151> 2001-08-03

<150> US 60/311,551
<151> 2001-08-10

<150> US 60/363,649
<151> 2002-03-08

<160> 112

<170> PERL Program

<210> 1
<211> 246
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2867236CD1

<400> 1
Met Val Val Asp Phe Trp Thr Trp Glu Gln Thr Phe Gln Glu Leu
1 5 10 15
Ile Gln Glu Ala Lys Pro Arg Ala Thr Trp Thr Leu Lys Leu Asp
20 25 30
Gly Asn Leu Gln Leu Asp Cys Leu Ala Gln Gly Trp Lys Gln Tyr
35 40 45
Gln Gln Arg Ala Phe Gly Trp Phe Arg Cys Ser Ser Cys Gln Arg
50 55 60
Ser Trp Ala Ser Ala Gln Val Gln Ile Leu Cys His Thr Tyr Trp
65 70 75
Glu His Trp Thr Ser Gln Gly Gln Val Arg Met Arg Leu Phe Gly
80 85 90
Gln Arg Cys Gln Lys Cys Ser Trp Ser Gln Tyr Glu Met Pro Glu
95 100 105
Phe Ser Ser Asp Ser Thr Met Arg Ile Leu Ser Asn Leu Val Gln
110 115 120
His Ile Leu Lys Lys Tyr Tyr Gly Asn Gly Thr Arg Lys Ser Pro
125 130 135
Glu Met Pro Val Ile Leu Glu Val Ser Leu Glu Gly Ser His Asp
140 145 150
Thr Ala Asn Cys Glu Ala Cys Thr Leu Gly Ile Cys Gly Gln Gly
155 160 165
Leu Lys Ser Tyr Met Thr Lys Pro Ser Lys Ser Leu Leu Pro His
170 175 180
Leu Lys Thr Gly Asn Ser Ser Pro Gly Ile Gly Ala Val Tyr Leu
185 190 195
Ala Asn Gln Ala Lys Asn Gln Ser Ala Glu Ala Lys Glu Ala Lys
200 205 210
Gly Ser Gly Tyr Glu Lys Leu Gly Pro Ser Arg Asp Pro Asp Pro
215 220 225
Leu Asn Ile Cys Val Phe Ile Leu Leu Leu Val Phe Ile Val Val
230 235 240
Lys Cys Phe Thr Ser Glu
245

<210> 2
<211> 325

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1294096CD1

<400> 2
```
Met Ala Leu Ala Asp Ser Thr Arg Gly Leu Pro Asn Gly Gly Gly
 1               5                  10                  15
Gly Gly Gly Gly Ser Gly Ser Ser Ser Ser Ser Ala Glu Pro Pro
                20                  25                  30
Leu Phe Pro Asp Ile Val Glu Leu Asn Val Gly Gly Gln Val Tyr
                35                  40                  45
Val Thr Arg Arg Cys Thr Val Val Ser Val Pro Asp Ser Leu Leu
                50                  55                  60
Trp Arg Met Phe Thr Gln Gln Gln Pro Gln Glu Leu Ala Arg Asp
                65                  70                  75
Ser Lys Gly Arg Phe Phe Leu Asp Arg Asp Gly Phe Leu Phe Arg
                80                  85                  90
Tyr Ile Leu Asp Tyr Leu Arg Asp Leu Gln Leu Val Leu Pro Asp
                95                 100                 105
Tyr Phe Pro Glu Arg Ser Arg Leu Gln Arg Glu Ala Glu Tyr Phe
               110                 115                 120
Glu Leu Pro Glu Leu Val Arg Arg Leu Gly Ala Pro Gln Gln Pro
               125                 130                 135
Gly Pro Gly Pro Pro Pro Ser Arg Arg Gly Val His Lys Glu Gly
               140                 145                 150
Ser Leu Gly Asp Glu Leu Leu Pro Leu Gly Tyr Ser Glu Pro Glu
               155                 160                 165
Gln Gln Glu Gly Ala Ser Ala Gly Ala Pro Ser Pro Thr Leu Glu
               170                 175                 180
Leu Ala Ser Arg Ser Pro Ser Gly Gly Ala Ala Gly Pro Leu Leu
               185                 190                 195
Thr Pro Ser Gln Ser Leu Asp Gly Ser Arg Arg Ser Gly Tyr Ile
               200                 205                 210
Thr Ile Gly Tyr Arg Gly Ser Tyr Thr Ile Gly Arg Asp Ala Gln
               215                 220                 225
Ala Asp Ala Lys Phe Arg Arg Val Ala Arg Ile Thr Val Cys Gly
               230                 235                 240
Lys Thr Ser Leu Ala Lys Glu Val Phe Gly Asp Thr Leu Asn Glu
               245                 250                 255
Ser Arg Asp Pro Asp Arg Pro Pro Glu Arg Tyr Thr Ser Arg Tyr
               260                 265                 270
Tyr Leu Lys Phe Asn Phe Leu Glu Gln Ala Phe Asp Lys Leu Ser
               275                 280                 285
Glu Ser Gly Phe His Met Val Ala Cys Ser Ser Thr Gly Thr Cys
               290                 295                 300
Ala Phe Ala Ser Ser Thr Asp Gln Ser Glu Asp Lys Ile Trp Thr
               305                 310                 315
Ser Tyr Thr Glu Tyr Val Phe Cys Arg Glu
               320                 325
```

<210> 3
<211> 376
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7238537CD1

<400> 3

Met Ala Arg Gly Pro Gly Pro Leu Gly Arg Pro Arg Pro Asp Thr
1 5 10 15
Val Ala Met Pro Lys Arg Gly Lys Arg Leu Lys Phe Arg Ala His
20 25 30
Asp Ala Cys Ser Gly Arg Val Thr Val Ala Asp Tyr Ala Asn Ser
35 40 45
Asp Pro Ala Val Val Arg Ser Gly Arg Val Lys Lys Ala Val Ala
50 55 60
Asn Ala Val Gln Gln Glu Val Lys Ser Leu Cys Gly Leu Glu Ala
65 70 75
Ser Gln Val Pro Ala Glu Glu Ala Leu Ser Gly Ala Gly Glu Pro
80 85 90
Cys Asp Ile Ile Asp Ser Ser Asp Glu Met Asp Ala Gln Glu Glu
95 100 105
Ser Ile His Glu Arg Thr Val Ser Arg Lys Lys Lys Ser Lys Arg
110 115 120
His Lys Glu Glu Leu Asp Gly Ala Gly Gly Glu Glu Tyr Pro Met
125 130 135
Asp Ile Trp Leu Leu Leu Ala Ser Tyr Ile Arg Pro Glu Asp Ile
140 145 150
Val Asn Phe Ser Leu Ile Cys Lys Asn Ala Trp Thr Val Thr Cys
155 160 165
Thr Ala Ala Phe Trp Thr Arg Leu Tyr Arg Arg His Tyr Thr Leu
170 175 180
Asp Ala Ser Leu Pro Leu Arg Leu Arg Pro Glu Ser Met Glu Lys
185 190 195
Leu Arg Cys Leu Arg Ala Cys Val Ile Arg Ser Leu Tyr His Met
200 205 210
Tyr Glu Pro Phe Ala Ala Arg Ile Ser Lys Asn Pro Ala Ile Pro
215 220 225
Glu Ser Thr Pro Ser Thr Leu Lys Asn Ser Lys Cys Leu Leu Phe
230 235 240
Trp Cys Arg Lys Ile Val Gly Asn Arg Gln Glu Pro Met Trp Glu
245 250 255
Phe Asn Phe Lys Phe Lys Lys Gln Ser Pro Arg Leu Lys Ser Lys
260 265 270
Cys Thr Gly Gly Leu Gln Pro Pro Val Gln Tyr Glu Asp Val His
275 280 285
Thr Asn Pro Asp Gln Asp Cys Cys Leu Leu Gln Val Thr Thr Leu
290 295 300
Asn Phe Ile Phe Ile Pro Ile Val Met Gly Met Ile Phe Thr Leu
305 310 315
Phe Thr Ile Asn Val Ser Thr Asp Met Arg His His Arg Val Arg
320 325 330
Leu Val Phe Gln Asp Ser Pro Val His Gly Gly Arg Lys Leu Arg
335 340 345
Ser Glu Gln Gly Val Gln Val Ile Leu Asp Pro Val His Ser Val
350 355 360
Arg Leu Phe Asp Trp Trp His Pro Gln Tyr Pro Phe Ser Leu Arg
365 370 375
Ala

<210> 4
<211> 461
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7494391CD1

<400> 4

```
Met Lys Ile Leu Phe Val Glu Pro Ala Ile Phe Leu Ser Ala Phe
 1               5                  10                  15
Ala Met Thr Leu Thr Gly Pro Leu Thr Thr Gln Tyr Val Tyr Arg
                20                  25                  30
Arg Ile Trp Glu Glu Thr Gly Asn Tyr Thr Phe Ser Ser Asp Ser
                35                  40                  45
Asn Ile Ser Glu Cys Glu Lys Asn Lys Ser Ser Pro Ile Phe Ala
                50                  55                  60
Phe Gln Glu Glu Val Gln Lys Lys Val Ser Arg Phe Asn Leu Gln
                65                  70                  75
Met Asp Ile Ser Gly Leu Ile Pro Gly Leu Val Ser Thr Phe Ile
                80                  85                  90
Leu Leu Ser Ile Ser Asp His Tyr Gly Arg Lys Phe Pro Met Ile
                95                  100                 105
Leu Ser Ser Val Gly Ala Leu Ala Thr Ser Val Trp Leu Cys Leu
                110                 115                 120
Leu Cys Tyr Phe Ala Phe Pro Phe Gln Leu Leu Ile Ala Ser Thr
                125                 130                 135
Phe Ile Gly Ala Phe Cys Gly Asn Tyr Thr Thr Phe Trp Gly Ala
                140                 145                 150
Cys Phe Ala Tyr Ile Val Asp Gln Cys Lys Glu His Lys Gln Lys
                155                 160                 165
Thr Ile Arg Ile Ala Ile Ile Asp Phe Leu Leu Gly Leu Val Thr
                170                 175                 180
Gly Leu Thr Gly Leu Ser Ser Gly Tyr Phe Ile Arg Glu Leu Gly
                185                 190                 195
Phe Glu Trp Ser Phe Leu Ile Ile Ala Val Ser Leu Ala Val Asn
                200                 205                 210
Leu Ile Tyr Ile Leu Phe Phe Leu Gly Asp Pro Val Lys Glu Cys
                215                 220                 225
Ser Ser Gln Asn Val Thr Met Ser Cys Ser Glu Gly Phe Lys Asn
                230                 235                 240
Leu Phe Tyr Arg Thr Tyr Met Leu Phe Lys Asn Ala Ser Gly Lys
                245                 250                 255
Arg Arg Phe Leu Leu Cys Leu Leu Leu Phe Thr Val Ile Thr Tyr
                260                 265                 270
Phe Phe Val Val Ile Gly Ile Ala Pro Ile Phe Ile Leu Tyr Glu
                275                 280                 285
Leu Asp Ser Pro Leu Cys Trp Asn Glu Val Phe Ile Gly Tyr Gly
                290                 295                 300
Ser Ala Leu Gly Ser Ala Ser Phe Leu Thr Ser Phe Leu Gly Ile
                305                 310                 315
Trp Leu Phe Ser Tyr Cys Met Glu Asp Ile His Met Ala Phe Ile
                320                 325                 330
Gly Ile Phe Thr Thr Met Thr Gly Met Ala Met Thr Ala Phe Ala
                335                 340                 345
Ser Thr Thr Leu Met Met Phe Leu Ala Arg Val Pro Phe Leu Phe
                350                 355                 360
Thr Ile Val Pro Phe Ser Val Leu Arg Ser Met Leu Ser Lys Val
                365                 370                 375
Val Arg Ser Thr Glu Gln Gly Thr Leu Phe Ala Cys Ile Ala Phe
                380                 385                 390
Leu Glu Thr Leu Gly Gly Val Thr Ala Val Ser Thr Phe Asn Gly
                395                 400                 405
Ile Tyr Ser Ala Thr Val Ala Trp Tyr Pro Gly Phe Thr Phe Leu
                410                 415                 420
Leu Ser Ala Gly Leu Leu Leu Leu Pro Ala Ile Ser Leu Cys Val
                425                 430                 435
Val Lys Cys Thr Ser Trp Asn Glu Gly Ser Tyr Glu Leu Leu Ile
                440                 445                 450
Gln Glu Glu Ser Ser Glu Asp Ala Ser Asp Arg
                455                 460
```

<210> 5
<211> 168
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6451054CD1

<400> 5
Met Met Glu Glu Ile Asp Arg Phe Gln Val Pro Thr Ala His Ser
1 5 10 15
Glu Met Gln Pro Leu Asp Pro Ala Ala Ala Ser Ile Ser Asp Gly
20 25 30
Asp Cys Asp Ala Arg Glu Glu Lys Gln Arg Glu Leu Ala Arg Lys
35 40 45
Gly Ser Leu Lys Asn Gly Ser Met Gly Ser Pro Val Asn Gln Gln
50 55 60
Pro Lys Lys Asn Asn Val Met Ala Arg Thr Arg Leu Val Val Pro
65 70 75
Asn Lys Gly Tyr Ser Ser Leu Asp Gln Ser Pro Asp Glu Lys Pro
80 85 90
Leu Val Ala Leu Asp Thr Asp Ser Asp Asp Asp Phe Asp Met Ser
95 100 105
Arg Tyr Ser Ser Ser Gly Tyr Ser Ser Ala Glu Gln Ile Asn Gln
110 115 120
Asp Leu Asn Ile Gln Leu Leu Lys Asp Gly Tyr Arg Leu Asp Glu
125 130 135
Ile Pro Asp Asp Glu Asp Leu Asp Leu Ile Pro Pro Lys Ser Val
140 145 150
Asn Pro Thr Cys Met Cys Cys Gln Ala Thr Ser Ser Thr Ala Cys
155 160 165
His Ile Gln

<210> 6
<211> 832
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7494592CD1

<400> 6
Met Met Glu Glu Glu Glu Leu Glu Phe Val Glu Glu Leu Glu Ala
1 5 10 15
Val Leu Gln Leu Thr Pro Glu Val Gln Leu Ala Ile Glu Gln Val
20 25 30
Phe Pro Ser Gln Asp Pro Leu Asp Arg Ala Asp Phe Asn Ala Val
35 40 45
Glu Tyr Ile Asn Thr Leu Phe Pro Thr Glu Gln Ser Leu Ala Asn
50 55 60
Ile Asp Glu Val Val Asn Lys Ile Arg Leu Lys Ile Arg Arg Leu
65 70 75
Asp Asp Asn Ile Arg Thr Val Val Arg Gly Gln Thr Asn Val Gly
80 85 90
Gln Asp Gly Arg Gln Ala Leu Glu Glu Ala Gln Lys Ala Ile Gln
95 100 105
Gln Leu Phe Gly Lys Ile Lys Asp Ile Lys Asp Lys Ala Glu Lys
110 115 120
Ser Glu Gln Met Val Lys Glu Ile Thr Arg Asp Ile Lys Gln Leu
125 130 135

```
Asp His Ala Lys Arg His Leu Thr Thr Ser Ile Thr Thr Leu Asn
                140                 145                 150
His Leu His Met Leu Ala Gly Gly Val Asp Ser Leu Glu Ala Met
                155                 160                 165
Thr Arg Arg Arg Gln Tyr Gly Glu Val Ala Asn Leu Leu Gln Gly
                170                 175                 180
Val Met Asn Val Leu Glu His Phe His Lys Tyr Met Gly Ile Pro
                185                 190                 195
Gln Ile Arg Gln Leu Ser Glu Arg Val Lys Ala Ala Gln Thr Glu
                200                 205                 210
Leu Gly Gln Gln Ile Leu Ala Asp Phe Glu Glu Ala Phe Pro Ser
                215                 220                 225
Gln Gly Thr Lys Arg Pro Gly Gly Pro Ser Asn Val Leu Arg Asp
                230                 235                 240
Ala Cys Leu Val Ala Asn Ile Leu Asp Pro Arg Ile Lys Gln Glu
                245                 250                 255
Ile Ile Lys Lys Phe Ile Lys Gln His Leu Ser Glu Tyr Leu Val
                260                 265                 270
Leu Phe Gln Glu Asn Gln Asp Val Ala Trp Leu Asp Lys Ile Asp
                275                 280                 285
Arg Arg Tyr Ala Trp Ile Lys Arg Gln Leu Val Asp Tyr Glu Glu
                290                 295                 300
Lys Tyr Gly Arg Met Phe Pro Arg Glu Trp Cys Met Ala Glu Arg
                305                 310                 315
Ile Ala Val Glu Phe Cys His Val Thr Arg Ala Glu Leu Ala Lys
                320                 325                 330
Ile Met Arg Thr Arg Ala Lys Glu Ile Glu Val Lys Leu Leu Leu
                335                 340                 345
Phe Ala Ile Gln Arg Thr Thr Asn Phe Glu Gly Phe Leu Ala Lys
                350                 355                 360
Arg Phe Ser Gly Cys Thr Leu Thr Asp Gly Thr Leu Lys Lys Leu
                365                 370                 375
Glu Ser Pro Pro Pro Ser Thr Asn Pro Phe Leu Glu Asp Glu Pro
                380                 385                 390
Thr Pro Glu Met Glu Glu Leu Ala Thr Glu Lys Gly Asp Leu Asp
                395                 400                 405
Gln Pro Lys Lys Pro Lys Ala Pro Asp Asn Pro Phe His Gly Ile
                410                 415                 420
Val Ser Lys Cys Phe Glu Pro His Leu Tyr Val Tyr Ile Glu Ser
                425                 430                 435
Gln Asp Lys Asn Leu Gly Glu Leu Ile Asp Arg Phe Val Ala Asp
                440                 445                 450
Phe Lys Ala Gln Gly Pro Pro Lys Pro Asn Thr Asp Glu Gly Gly
                455                 460                 465
Ala Val Leu Pro Ser Cys Ala Asp Leu Phe Val Tyr Tyr Lys Lys
                470                 475                 480
Cys Met Val Gln Cys Ser Gln Leu Ser Thr Gly Glu Pro Met Ile
                485                 490                 495
Ala Leu Thr Thr Ile Phe Gln Lys Tyr Leu Arg Glu Tyr Ala Trp
                500                 505                 510
Lys Ile Leu Ser Gly Asn Leu Pro Lys Thr Thr Thr Ser Ser Gly
                515                 520                 525
Gly Leu Thr Ile Ser Ser Leu Leu Lys Glu Lys Glu Gly Ser Glu
                530                 535                 540
Val Ala Lys Phe Thr Leu Glu Glu Leu Cys Leu Ile Cys Asn Ile
                545                 550                 555
Leu Ser Thr Ala Glu Tyr Cys Leu Ala Thr Thr Gln Gln Leu Glu
                560                 565                 570
Glu Lys Leu Lys Glu Lys Val Asp Val Ser Leu Ile Glu Arg Ile
                575                 580                 585
Asn Leu Thr Gly Glu Met Asp Thr Phe Ser Thr Val Ile Ser Ser
                590                 595                 600
Ser Ile Gln Leu Leu Val Gln Asp Leu Asp Ala Ala Cys Asp Pro
```

```
                    605                     610                     615
Ala Leu Thr Ala Met Ser Lys Met Gln Trp Gln Asn Val Glu His
                    620                     625                     630
Val Gly Asp Gln Ser Pro Tyr Val Thr Ser Val Ile Leu His Ile
                    635                     640                     645
Lys Gln Asn Val Pro Ile Ile Arg Asp Asn Leu Ala Ser Thr Arg
                    650                     655                     660
Lys Tyr Phe Thr Gln Phe Cys Val Lys Phe Ala Asn Ser Phe Ile
                    665                     670                     675
Pro Lys Phe Ile Thr His Leu Phe Lys Cys Lys Pro Ile Ser Met
                    680                     685                     690
Val Gly Ala Glu Gln Leu Leu Leu Asp Thr His Ser Leu Lys Met
                    695                     700                     705
Val Leu Leu Asp Leu Pro Ser Ile Ser Ser Gln Val Val Arg Lys
                    710                     715                     720
Ala Pro Ala Ser Tyr Thr Lys Ile Val Val Lys Gly Met Thr Arg
                    725                     730                     735
Ala Glu Met Ile Leu Lys Val Val Met Ala Pro His Glu Pro Leu
                    740                     745                     750
Val Val Phe Val Asp Asn Tyr Ile Lys Leu Leu Thr Asp Cys Asn
                    755                     760                     765
Thr Glu Thr Phe Gln Lys Ile Leu Asp Met Lys Gly Leu Lys Arg
                    770                     775                     780
Ser Glu Gln Ser Ser Met Leu Glu Leu Leu Arg Gln Arg Leu Pro
                    785                     790                     795
Ala Pro Pro Ser Gly Ala Glu Ser Ser Gly Ser Leu Ser Leu Thr
                    800                     805                     810
Ala Pro Thr Pro Glu Gln Glu Ser Ser Arg Ile Arg Lys Leu Glu
                    815                     820                     825
Lys Leu Ile Lys Lys Arg Leu
                    830
```

<210> 7
<211> 393
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5202657CD1

<400> 7

```
Met Glu Gln Cys Ala Cys Val Glu Arg Glu Leu Asp Lys Val Leu
 1               5                  10                  15
Gln Lys Phe Leu Thr Tyr Gly Gln His Cys Glu Arg Ser Leu Glu
                20                  25                  30
Glu Leu Leu His Tyr Val Gly Gln Leu Arg Ala Glu Leu Ala Ser
                35                  40                  45
Ala Ala Leu Gln Gly Thr Pro Leu Ser Ala Thr Leu Ser Leu Val
                50                  55                  60
Met Ser Gln Cys Cys Arg Lys Ile Lys Asp Thr Val Gln Lys Leu
                65                  70                  75
Ala Ser Asp His Lys Asp Ile His Ser Ser Val Ser Arg Val Gly
                80                  85                  90
Lys Ala Ile Asp Arg Asn Phe Asp Ser Glu Ile Cys Gly Val Val
                95                 100                 105
Ser Asp Ala Val Trp Asp Ala Arg Glu Gln Gln Gln Gln Ile Leu
               110                 115                 120
Gln Met Ala Ile Val Glu His Leu Tyr Gln Gln Gly Met Leu Ser
               125                 130                 135
Val Ala Glu Glu Leu Cys Gln Glu Ser Thr Leu Asn Val Asp Leu
               140                 145                 150
Asp Phe Lys Gln Pro Phe Leu Glu Leu Asn Arg Ile Leu Glu Ala
```

```
                155                 160                 165
Leu His Glu Gln Asp Leu Gly Pro Ala Leu Glu Trp Ala Val Ser
                170                 175                 180
His Arg Gln Arg Leu Leu Glu Leu Asn Ser Ser Leu Glu Phe Lys
                185                 190                 195
Leu His Arg Leu His Phe Ile Arg Leu Leu Ala Gly Gly Pro Ala
                200                 205                 210
Lys Gln Leu Glu Ala Leu Ser Tyr Ala Arg His Phe Gln Pro Phe
                215                 220                 225
Ala Arg Leu His Gln Arg Glu Ile Gln Val Met Met Gly Ser Leu
                230                 235                 240
Val Tyr Leu Arg Leu Gly Leu Glu Lys Ser Pro Tyr Cys His Leu
                245                 250                 255
Leu Asp Ser Ser His Trp Ala Glu Ile Cys Glu Thr Phe Thr Arg
                260                 265                 270
Asp Ala Cys Ser Leu Leu Gly Leu Ser Val Glu Ser Pro Leu Ser
                275                 280                 285
Val Ser Phe Ala Ser Gly Cys Val Ala Leu Pro Val Leu Met Asn
                290                 295                 300
Ile Lys Ala Val Ile Glu Gln Arg Gln Cys Thr Gly Val Trp Asn
                305                 310                 315
His Lys Asp Glu Leu Pro Ile Glu Ile Glu Leu Gly Met Lys Cys
                320                 325                 330
Trp Tyr His Ser Val Phe Ala Cys Pro Ile Leu Arg Gln Gln Thr
                335                 340                 345
Ser Asp Ser Asn Pro Pro Ile Lys Leu Ile Cys Gly His Val Ile
                350                 355                 360
Ser Arg Asp Ala Leu Asn Lys Leu Ile Asn Gly Gly Lys Leu Lys
                365                 370                 375
Cys Pro Tyr Cys Pro Met Glu Gln Asn Pro Ala Asp Gly Lys Arg
                380                 385                 390
Ile Ile Phe


<210> 8
<211> 280
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2013529CD1

<400> 8
Met Ala Thr Glu Ala Pro Val Asn Ile Ala Pro Pro Glu Cys Ser
 1               5                   10                  15
Thr Val Val Ser Thr Ala Val Asp Ser Leu Ile Trp Gln Pro Asn
                20                  25                  30
Ser Leu Asn Met His Met Ile Arg Pro Lys Ser Ala Lys Gly Arg
                35                  40                  45
Thr Arg Pro Ser Leu Gln Lys Ser Gln Gly Val Glu Val Cys Ala
                50                  55                  60
His His Ile Pro Ser Pro Pro Pro Ala Ile Pro Tyr Glu Leu Pro
                65                  70                  75
Ser Ser Gln Lys Pro Gly Ala Cys Ala Pro Lys Ser Pro Asn Gln
                80                  85                  90
Gly Ala Ser Asp Glu Ile Pro Glu Leu Gln Gln Gln Val Pro Thr
                95                  100                 105
Gly Ala Ser Ser Ser Leu Asn Lys Tyr Pro Val Leu Pro Ser Ile
                110                 115                 120
Asn Arg Lys Asn Leu Glu Glu Glu Ala Val Glu Thr Val Ala Lys
                125                 130                 135
Lys Ala Ser Ser Leu Gln Leu Ser Ser Ile Arg Ala Leu Tyr Gln
```

```
                140                 145                 150
Asp Glu Thr Gly Thr Met Lys Thr Ser Glu Glu Asp Ser Arg Ala
                155                 160                 165
Arg Ala Cys Ala Val Glu Arg Lys Phe Ile Val Arg Thr Lys Lys
                170                 175                 180
Gln Gly Ser Ser Arg Ala Gly Asn Leu Glu Glu Pro Ser Asp Gln
                185                 190                 195
Glu Pro Arg Leu Leu Leu Ala Val Arg Ser Pro Thr Gly Gln Arg
                200                 205                 210
Phe Val Arg His Phe Arg Pro Thr Asp Asp Leu Gln Thr Ile Val
                215                 220                 225
Ala Val Ala Glu Gln Lys Asn Lys Thr Ser Tyr Arg His Cys Ser
                230                 235                 240
Ile Glu Thr Met Glu Val Pro Arg Arg Arg Phe Ser Asp Leu Thr
                245                 250                 255
Lys Ser Leu Gln Glu Cys Arg Ile Pro His Lys Ser Val Leu Gly
                260                 265                 270
Ile Ser Leu Glu Asp Gly Glu Gly Trp Pro
                275                 280
```

<210> 9
<211> 344
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3841351CD1

<400> 9

```
Met Asp Ser Tyr Ser Ala Pro Glu Ser Thr Pro Ser Ala Ser Ser
  1               5                  10                  15
Arg Pro Glu Asp Tyr Phe Ile Gly Ala Thr Pro Leu Gln Lys Arg
                 20                  25                  30
Leu Glu Ser Val Arg Lys Gln Ser Ser Phe Ile Leu Thr Pro Pro
                 35                  40                  45
Arg Arg Lys Ile Pro Gln Cys Ser Gln Leu Gln Glu Asp Val Asp
                 50                  55                  60
Pro Gln Lys Val Ala Phe Leu Leu His Lys Gln Trp Thr Leu Tyr
                 65                  70                  75
Ser Leu Thr Pro Leu Tyr Lys Phe Ser Tyr Ser Asn Leu Lys Glu
                 80                  85                  90
Tyr Ser Arg Leu Leu Asn Ala Phe Ile Val Ala Glu Lys Gln Lys
                 95                 100                 105
Gly Leu Ala Val Glu Val Gly Glu Asp Phe Asn Ile Lys Val Ile
                110                 115                 120
Phe Ser Thr Leu Leu Gly Met Lys Gly Thr Gln Arg Asp Pro Glu
                125                 130                 135
Ala Phe Leu Val Gln Ile Val Ser Lys Ser Gln Leu Pro Ser Glu
                140                 145                 150
Asn Arg Glu Gly Lys Val Leu Trp Thr Gly Trp Phe Cys Cys Val
                155                 160                 165
Phe Gly Asp Ser Leu Leu Glu Thr Val Ser Glu Asp Phe Thr Cys
                170                 175                 180
Leu Pro Leu Phe Leu Ala Asn Gly Ala Glu Ser Asn Thr Ala Ile
                185                 190                 195
Ile Gly Thr Trp Phe Gln Lys Thr Phe Asp Cys Tyr Phe Ser Pro
                200                 205                 210
Leu Ala Ile Asn Ala Phe Asn Leu Ser Trp Met Ala Ala Met Trp
                215                 220                 225
Thr Ala Cys Lys Met Asp His Tyr Val Ala Thr Thr Glu Phe Leu
                230                 235                 240
Trp Ser Val Pro Cys Ser Pro Gln Ser Leu Asp Ile Ser Phe Ala
```

```
                245                 250                 255
Ile His Pro Glu Asp Ala Lys Ala Leu Trp Asp Ser Val His Lys
                260                 265                 270
Thr Pro Gly Glu Val Thr Gln Glu Glu Val Asp Leu Phe Met Asp
                275                 280                 285
Cys Leu Tyr Ser His Phe His Arg His Phe Lys Ile His Leu Ser
                290                 295                 300
Ala Thr Arg Leu Val Arg Val Ser Thr Ser Val Ala Ser Ala His
                305                 310                 315
Thr Asp Gly Lys Ile Lys Ile Leu Cys His Lys Tyr Leu Ile Gly
                320                 325                 330
Val Leu Ala Tyr Leu Thr Glu Leu Ala Ile Phe Gln Ile Glu
                335                 340

<210> 10
<211> 405
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 152116CD1

<400> 10
Met Glu Pro Gly Ala Gly Gly Arg Asn Thr Ala Arg Ala Gln Arg
 1               5                  10                  15
Ala Gly Ser Pro Asn Thr Pro Pro Pro Arg Glu Gln Glu Arg Lys
                20                  25                  30
Leu Glu Gln Glu Lys Leu Ser Gly Val Val Lys Ser Val His Arg
                35                  40                  45
Arg Leu Arg Lys Lys Tyr Arg Glu Val Gly Asp Phe Asp Lys Ile
                50                  55                  60
Trp Arg Glu His Cys Glu Asp Glu Glu Thr Leu Cys Glu Tyr Ala
                65                  70                  75
Val Ala Met Lys Asn Leu Ala Asp Asn His Trp Ala Lys Thr Cys
                80                  85                  90
Glu Gly Glu Gly Arg Ile Glu Trp Cys Cys Ser Val Cys Arg Glu
                95                  100                 105
Tyr Phe Gln Asn Gly Gly Lys Arg Lys Ala Leu Glu Lys Asp Glu
                110                 115                 120
Lys Arg Ala Val Leu Ala Thr Lys Thr Thr Pro Ala Leu Asn Met
                125                 130                 135
His Glu Ser Ser Gln Leu Glu Gly His Leu Thr Asn Leu Ser Phe
                140                 145                 150
Thr Asn Pro Glu Phe Ile Thr Glu Leu Leu Gln Ala Ser Gly Lys
                155                 160                 165
Ile Arg Leu Leu Asp Val Gly Ser Cys Phe Asn Pro Phe Leu Lys
                170                 175                 180
Phe Glu Glu Phe Leu Thr Val Gly Ile Asp Ile Val Pro Ala Val
                185                 190                 195
Glu Ser Val Tyr Lys Cys Asp Phe Leu Asn Leu Gln Leu Gln Gln
                200                 205                 210
Pro Leu Gln Leu Ala Gln Asp Ala Ile Asp Ala Phe Leu Lys Gln
                215                 220                 225
Leu Lys Asn Pro Ile Asp Ser Leu Pro Gly Glu Leu Phe His Val
                230                 235                 240
Val Val Phe Ser Leu Leu Leu Ser Tyr Phe Pro Ser Pro Tyr Gln
                245                 250                 255
Arg Trp Ile Cys Cys Lys Lys Ala His Glu Leu Leu Val Leu Asn
                260                 265                 270
Gly Leu Leu Leu Ile Ile Thr Pro Asp Ser Ser His Gln Asn Arg
                275                 280                 285
His Ala Met Met Met Lys Ser Trp Lys Ile Ala Ile Glu Ser Leu
```

```
                290                 295                 300
Gly Phe Lys Arg Phe Lys Tyr Ser Lys Phe Ser His Met His Leu
                305                 310                 315
Met Ala Phe Arg Lys Ile Ser Leu Lys Thr Thr Ser Asp Leu Val
                320                 325                 330
Ser Arg Asn Tyr Pro Gly Met Leu Tyr Ile Pro Gln Asp Phe Asn
                335                 340                 345
Ser Ile Glu Asp Glu Glu Tyr Ser Asn Pro Ser Cys Tyr Val Arg
                350                 355                 360
Ser Asp Ile Glu Asp Glu Gln Leu Ala Tyr Gly Phe Thr Glu Leu
                365                 370                 375
Pro Asp Ala Pro Tyr Asp Ser Asp Ser Gly Glu Ser Gln Ala Ser
                380                 385                 390
Ser Ile Pro Phe Tyr Glu Leu Glu Asp Pro Ile Leu Leu Leu Ser
                395                 400                 405


<210> 11
<211> 185
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2381031CD1

<400> 11
Met Glu Val His Gly Lys Pro Lys Ala Ser Pro Ser Cys Ser Ser
  1             5                  10                  15
Pro Thr Arg Asp Ser Ser Gly Val Pro Val Ser Lys Glu Leu Leu
                20                 25                  30
Thr Ala Gly Ser Asp Gly Arg Gly Gly Ile Trp Asp Arg Leu Leu
                35                 40                  45
Ile Asn Ser Gln Pro Lys Ser Arg Lys Thr Ser Thr Leu Gln Thr
                50                 55                  60
Val Arg Ile Glu Arg Ser Pro Leu Leu Asp Gln Val Gln Thr Phe
                65                 70                  75
Leu Pro Gln Met Ala Arg Ala Asn Glu Lys Leu Arg Lys Glu Met
                80                 85                  90
Ala Ala Ala Pro Pro Gly Arg Phe Asn Ile Glu Asn Ile Asp Gly
                95                 100                 105
Pro His Ser Lys Val Ile Gln Met Asp Val Ala Leu Phe Glu Met
                110                115                 120
Asn Gln Ser Asp Ser Lys Glu Val Asp Ser Ser Glu Glu Ser Ser
                125                130                 135
Gln Asp Ser Ser Glu Asn Ser Ser Glu Ser Glu Asp Glu Asp Asp
                140                145                 150
Ser Ile Pro Ser Glu Val Thr Ile Asp Asn Ile Lys Leu Pro Asn
                155                160                 165
Ser Glu Gly Gly Lys Gly Lys Ile Glu Val Leu Asp Ser Pro Ala
                170                175                 180
Ser Lys Lys Lys Lys
                185

<210> 12
<211> 463
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2511371CD1
```

<400> 12

```
Met Ala Gln Gln Gln Thr Gly Ser Arg Lys Arg Lys Ala Pro Ala
1               5                   10                  15
Val Glu Ala Asp Ala Glu Ser Ser Pro Ser Gln Gly Leu Ala Ala
                20                  25                  30
Ala Asp Gly Glu Gly Pro Leu Leu Leu Lys Arg Gln Arg Arg Pro
                35                  40                  45
Ala Thr Tyr Arg Ser Met Ala His Tyr Leu Lys Val Arg Glu Val
                50                  55                  60
Gly Gly Trp Gly Pro Ala Arg Leu Gln Gly Phe Asp Gly Glu Leu
                65                  70                  75
Arg Gly Tyr Ala Val Gln Arg Leu Pro Glu Leu Leu Thr Glu Arg
                80                  85                  90
Gln Leu Glu Leu Gly Thr Val Asn Lys Val Phe Ala Ser Gln Trp
                95                  100                 105
Leu Asn Ser Arg Gln Val Val Cys Gly Thr Lys Cys Asn Thr Leu
                110                 115                 120
Phe Val Val Asp Val Glu Ser Gly His Ile Ala Arg Ile Pro Leu
                125                 130                 135
Leu Arg Asp Ser Glu Ala Arg Leu Ala Gln Asp Gln Gln Gly Cys
                140                 145                 150
Gly Ile His Ala Ile Glu Leu Asn Pro Ser Lys Thr Leu Leu Ala
                155                 160                 165
Thr Gly Gly Glu Asn Pro Asn Ser Leu Ala Ile Tyr Gln Leu Pro
                170                 175                 180
Ser Leu Asp Pro Leu Cys Leu Gly Asp Arg His Gly His Lys Asp
                185                 190                 195
Trp Ile Phe Ala Val Ala Trp Leu Ser Asp Thr Val Ala Val Ser
                200                 205                 210
Gly Ser Arg Asp Gly Thr Val Ala Leu Trp Arg Met Asp Pro Asp
                215                 220                 225
Lys Phe Asp Asp Thr Val Ala Trp His Ser Glu Val Gly Leu Pro
                230                 235                 240
Val Tyr Ala His Ile Arg Pro Arg Asp Val Glu Ala Ile Pro Arg
                245                 250                 255
Ala Ile Ile Asn Pro Ser Asn Arg Lys Val Arg Ala Leu Ala Cys
                260                 265                 270
Gly Gly Lys Asn Gln Glu Leu Gly Ala Val Ser Leu Asp Gly Tyr
                275                 280                 285
Phe His Leu Trp Lys Ala Gly Ser Ala Leu Ser Arg Leu Leu Ser
                290                 295                 300
Ile Arg Leu Pro Tyr Phe Arg Asp Asn Val Cys Leu Thr Tyr Cys
                305                 310                 315
Asp Asp Met Ser Val Tyr Ala Val Gly Ser His Ser His Val Ser
                320                 325                 330
Phe Leu Asp Leu Arg Gln Asp Gln Gln Asn Ile Arg Pro Leu Cys
                335                 340                 345
Ser Arg Glu Gly Gly Thr Gly Val Arg Ser Leu Ser Phe Tyr Arg
                350                 355                 360
His Ile Ile Thr Val Gly Thr Gly Gln Gly Ser Leu Leu Phe Tyr
                365                 370                 375
Asp Val Arg Ala Gln Lys Phe Leu Glu Glu Arg Ala Ser Ala Thr
                380                 385                 390
Leu Glu Ser Ser Ser Gly Pro Ala Arg Arg Lys Leu Arg Leu Ala
                395                 400                 405
Cys Gly Arg Gly Trp Leu Asn His Asn Asp Phe Trp Val Asn Tyr
                410                 415                 420
Phe Gly Gly Met Glu Val Phe Pro Asn Ala Leu Tyr Thr His Cys
                425                 430                 435
Tyr Asn Trp Pro Glu Met Lys Leu Phe Val Ala Gly Gly Pro Leu
                440                 445                 450
Pro Ala Gly Leu His Gly Asn Tyr Ala Gly Leu Trp Ser
                455                 460
```

<210> 13
<211> 403
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8068623CD1

<400> 13
Met Ala Leu Arg Ser Ala Gln Gly Asp Gly Pro Thr Ser Gly His
1 5 10 15
Trp Asp Gly Gly Ala Glu Lys Ala Asp Phe Asn Ala Lys Arg Lys
20 25 30
Lys Lys Val Ala Glu Ile His Gln Ala Leu Asn Ser Asp Pro Thr
35 40 45
Asp Val Ala Ala Leu Arg Arg Met Ala Ile Ser Glu Gly Gly Leu
50 55 60
Leu Thr Asp Glu Ile Arg Arg Lys Val Trp Pro Lys Leu Leu Asn
65 70 75
Val Asn Ala Asn Asp Pro Pro Pro Ile Ser Gly Lys Asn Leu Arg
80 85 90
Gln Met Ser Lys Asp Tyr Gln Gln Val Leu Leu Asp Val Arg Arg
95 100 105
Ser Leu Arg Arg Phe Pro Pro Gly Met Pro Glu Glu Gln Arg Glu
110 115 120
Gly Leu Gln Glu Glu Leu Ile Asp Ile Ile Leu Leu Ile Leu Glu
125 130 135
Arg Asn Pro Gln Leu His Tyr Tyr Gln Gly Tyr His Asp Ile Val
140 145 150
Val Thr Phe Leu Leu Val Val Gly Glu Arg Leu Ala Thr Ser Leu
155 160 165
Val Glu Lys Leu Ser Thr His His Leu Arg Asp Phe Met Asp Pro
170 175 180
Thr Met Asp Asn Thr Lys His Ile Leu Asn Tyr Leu Met Pro Ile
185 190 195
Ile Asp Gln Val Asn Pro Glu Leu His Asp Phe Met Gln Ser Ala
200 205 210
Glu Val Gly Thr Ile Phe Ala Leu Ser Trp Leu Ile Thr Trp Phe
215 220 225
Gly His Val Leu Ser Asp Phe Arg His Val Val Arg Leu Tyr Asp
230 235 240
Phe Phe Leu Ala Cys His Pro Leu Met Pro Ile Tyr Phe Ala Ala
245 250 255
Val Ile Val Leu Tyr Arg Glu Gln Glu Val Leu Asp Cys Asp Cys
260 265 270
Asp Met Ala Ser Val His His Leu Leu Ser Gln Ile Pro Gln Asp
275 280 285
Leu Pro Tyr Glu Thr Leu Ile Ser Arg Ala Gly Asp Leu Phe Val
290 295 300
Gln Phe Pro Pro Ser Glu Leu Ala Arg Glu Ala Ala Ala Gln Gln
305 310 315
Gln Ala Glu Arg Thr Ala Ala Ser Thr Phe Lys Asp Phe Glu Leu
320 325 330
Ala Ser Ala Gln Gln Arg Pro Asp Met Val Leu Arg Gln Arg Phe
335 340 345
Arg Gly Leu Leu Arg Pro Glu Asp Arg Thr Lys Asp Val Leu Thr
350 355 360
Lys Pro Arg Thr Asn Arg Phe Val Lys Leu Ala Val Met Gly Leu
365 370 375
Thr Val Ala Leu Gly Ala Ala Ala Leu Ala Val Val Lys Ser Ala
380 385 390
Leu Glu Trp Ala Pro Lys Phe Gln Leu Gln Leu Phe Pro

395 400

<210> 14
<211> 574
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 677977CD1

<400> 14
Met Gly Gly Glu Arg Lys Ala Gln Thr Cys Ala Ala Thr Phe Ser
1 5 10 15
Val Pro Ala Arg Ala Cys Ala Ala Gly Ser Arg Thr Met Pro Thr
20 25 30
Cys Ala Gly Ser Trp Ser Ser Trp Ala Val Arg Trp Ala Leu Ser
35 40 45
Ala Arg Arg Cys Gly Trp Pro Thr Arg Arg Arg Ser Thr Ala Pro
50 55 60
Cys Thr Pro Gly Ser Trp Arg Cys Ala Thr Thr Gly Cys Arg Cys
65 70 75
Leu Ala Arg Ser Ser Arg Arg Ser Arg Gly Leu Arg Ala Pro Asp
80 85 90
Leu Arg Ala Arg Val His Leu Gln Gly Gln Pro Arg Leu Val Leu
95 100 105
Ala Leu Ala Glu Ala Pro Arg His Leu Gln Pro Ala Leu Leu Arg
110 115 120
Arg Gly Gly Pro Pro Ala Pro Ser Pro Ala Pro Gly Pro Pro Val
125 130 135
Lys Glu Glu Pro Ala Leu Pro Ser Gly Ala Gly Pro Leu Pro Asp
140 145 150
Arg Ala Pro Ala Pro Pro Pro Pro Ala Glu Gly Gly Tyr Gly Asp
155 160 165
Glu Gln Ile Tyr Ser Ala Ser Val Thr Gly Leu Tyr Trp Lys Leu
170 175 180
Leu Pro Glu Gln Ala Ala Pro Pro Gly Ala Gly Asp Pro Gly Ala
185 190 195
Gly Gly Cys Gly Arg Arg Trp Arg Gly Asp Arg Val Thr Val Leu
200 205 210
Leu Ala Ala Asn Leu Thr Gly Ser His Lys Leu Lys Pro Leu Val
215 220 225
Ile Gly Arg Leu Pro Asp Pro Pro Ser Leu Arg His His Asn Gln
230 235 240
Asp Lys Phe Pro Ala Ser Tyr Arg Tyr Ser Pro Asp Ala Trp Leu
245 250 255
Ser Arg Pro Leu Leu Arg Gly Trp Phe Phe Glu Glu Phe Val Pro
260 265 270
Gly Val Lys Arg Tyr Leu Arg Arg Ser Cys Leu Gln Gln Lys Ala
275 280 285
Val Leu Leu Val Ala His Pro Pro Cys Pro Ser Pro Ala Ala Ser
290 295 300
Met Pro Ala Leu Asp Ser Glu Asp Ala Pro Val Arg Cys Arg Pro
305 310 315
Glu Pro Leu Gly Pro Pro Glu Glu Leu Gln Thr Pro Asp Gly Ala
320 325 330
Val Arg Val Leu Phe Leu Ser Lys Gly Ser Ser Arg Ala His Ile
335 340 345
Pro Glu Pro Val Glu Gln Gly Val Val Ala Ala Phe Lys Gln Leu
350 355 360
Tyr Lys Arg Glu Leu Leu Arg Leu Ala Val Ser Cys Ala Ser Gly
365 370 375
Ser Pro Leu Asp Phe Met Arg Ser Phe Met Leu Lys Asp Met Leu

```
                380                 385                 390
Tyr Leu Ala Gly Leu Ser Trp Asp Leu Val Gln Ala Gly Ser Ile
                395                 400                 405
Glu Arg Cys Trp Leu Leu Gly Leu Arg Ala Ala Phe Glu Pro Arg
                410                 415                 420
Pro Gly Glu Asp Ser Ala Gly Gln Pro Ala Gln Ala Glu Glu Ala
                425                 430                 435
Ala Glu His Ser Arg Val Leu Ser Asp Leu Thr His Leu Ala Ala
                440                 445                 450
Leu Ala Tyr Lys Cys Leu Ala Pro Glu Glu Val Ala Glu Trp Leu
                455                 460                 465
His Leu Asp Asp Asp Gly Ala Ser Leu Pro Ser Ala Met Gly Gly
                470                 475                 480
Gly Glu Asp Glu Glu Glu Ala Thr Asp Tyr Gly Gly Thr Ser Ser
                485                 490                 495
Leu Pro Ser Ala Ile Gly Gly Gly Glu Asp Glu Glu Glu Ala Thr
                500                 505                 510
Asp Tyr Gly Gly Thr Ser Val Pro Thr Ala Gly Glu Ala Val Arg
                515                 520                 525
Gly Leu Glu Thr Ala Leu Arg Trp Leu Glu Asn Gln Asp Pro Arg
                530                 535                 540
Glu Val Gly Pro Leu Arg Leu Val Gln Leu Arg Ser Leu Ile Ser
                545                 550                 555
Met Ala Arg Arg Leu Gly Gly Ile Gly His Thr Pro Ala Gly Pro
                560                 565                 570
Tyr Asp Gly Val


&lt;210&gt; 15
&lt;211&gt; 731
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;220&gt;
&lt;221&gt; misc_feature
&lt;223&gt; Incyte ID No: 1661472CD1

&lt;400&gt; 15
Met Gln Gly Asp Pro Asp Asp Thr Ser His Arg Gly His Pro Leu
 1               5                  10                  15
Cys Lys Phe Cys Asp Glu Arg Tyr Leu Asp Asn Asp Glu Leu Leu
                20                  25                  30
Lys His Leu Arg Arg Asp His Tyr Phe Cys His Phe Cys Asp Ser
                35                  40                  45
Asp Gly Ala Gln Asp Tyr Tyr Ser Asp Tyr Ala Tyr Leu Arg Glu
                50                  55                  60
His Phe Arg Glu Lys His Phe Leu Cys Glu Glu Gly Arg Cys Ser
                65                  70                  75
Thr Glu Gln Phe Thr His Ala Phe Arg Thr Glu Ile Asp Leu Lys
                80                  85                  90
Ala His Arg Thr Ala Cys His Ser Arg Ser Arg Ala Glu Ala Arg
                95                  100                 105
Gln Asn Arg His Ile Asp Leu Gln Phe Ser Tyr Ala Pro Arg His
                110                 115                 120
Ser Arg Arg Asn Glu Gly Val Val Gly Gly Glu Asp Tyr Glu Glu
                125                 130                 135
Val Asp Arg Tyr Ser Arg Gln Gly Arg Val Ala Arg Ala Gly Thr
                140                 145                 150
Arg Gly Ala Gln Gln Ser Arg Arg Gly Ser Trp Arg Tyr Lys Arg
                155                 160                 165
Glu Glu Glu Asp Arg Glu Val Ala Ala Ala Val Arg Ala Ser Val
                170                 175                 180
Ala Ala Gln Gln Gln Glu Glu Ala Arg Arg Ser Glu Asp Gln Glu
```

185 190 195
Glu Gly Gly Arg Pro Lys Lys Glu Glu Ala Ala Ala Arg Gly Pro
200 205 210
Glu Asp Pro Arg Gly Pro Arg Arg Ser Pro Arg Thr Gln Gly Glu
215 220 225
Gly Pro Gly Pro Lys Glu Thr Ser Thr Asn Gly Pro Val Ser Gln
230 235 240
Glu Ala Phe Ser Val Thr Gly Pro Ala Ala Pro Gly Cys Val Gly
245 250 255
Val Pro Gly Ala Leu Pro Pro Pro Ser Pro Lys Leu Lys Asp Glu
260 265 270
Asp Phe Pro Ser Leu Ser Ala Ser Thr Ser Ser Ser Cys Ser Thr
275 280 285
Ala Ala Thr Pro Gly Pro Val Gly Leu Ala Leu Pro Tyr Ala Ile
290 295 300
Pro Ala Arg Gly Arg Ser Ala Phe Gln Glu Glu Asp Phe Pro Ala
305 310 315
Leu Val Ser Ser Val Pro Lys Pro Gly Thr Ala Pro Thr Ser Leu
320 325 330
Val Ser Ala Trp Asn Ser Ser Ser Ser Ser Lys Lys Val Ala Gln
335 340 345
Pro Pro Leu Ser Ala Gln Ala Thr Gly Ser Gly Gln Pro Thr Arg
350 355 360
Lys Ala Gly Lys Gly Ser Arg Gly Gly Arg Lys Gly Gly Pro Pro
365 370 375
Phe Thr Gln Glu Glu Glu Glu Asp Gly Gly Pro Ala Leu Gln Glu
380 385 390
Leu Leu Ser Thr Arg Pro Thr Gly Ser Val Ser Ser Thr Leu Gly
395 400 405
Leu Ala Ser Ile Gln Pro Ser Lys Val Gly Lys Lys Lys Lys Val
410 415 420
Gly Ser Glu Lys Pro Gly Thr Thr Leu Pro Gln Pro Pro Pro Ala
425 430 435
Thr Cys Pro Pro Gly Ala Leu Gln Ala Pro Glu Ala Pro Ala Ser
440 445 450
Arg Ala Glu Gly Pro Val Ala Val Val Val Asn Gly His Thr Glu
455 460 465
Gly Pro Ala Pro Ala Arg Ser Ala Pro Lys Glu Pro Pro Gly Leu
470 475 480
Pro Arg Pro Leu Gly Ser Phe Pro Cys Pro Thr Pro Gln Glu Asp
485 490 495
Phe Pro Ala Leu Gly Gly Pro Cys Pro Pro Arg Met Pro Pro Pro
500 505 510
Pro Gly Phe Ser Ala Val Val Leu Leu Lys Gly Thr Pro Pro Pro
515 520 525
Pro Pro Pro Gly Leu Val Pro Pro Ile Ser Lys Pro Pro Pro Gly
530 535 540
Phe Ser Gly Leu Leu Pro Ser Pro His Pro Ala Ser Val Pro Ser
545 550 555
Pro Ala Thr Thr Thr Thr Thr Lys Ala Pro Arg Leu Leu Pro Ala
560 565 570
Pro Arg Ala Tyr Leu Val Pro Glu Asn Phe Arg Glu Arg Asn Leu
575 580 585
Gln Leu Ile Gln Ser Ile Arg Asp Phe Leu Gln Ser Asp Glu Ala
590 595 600
Arg Phe Ser Glu Phe Lys Ser His Ser Gly Glu Phe Arg Gln Gly
605 610 615
Leu Ile Ser Ala Ala Gln Tyr Tyr Lys Ser Cys Arg Asp Leu Leu
620 625 630
Gly Glu Asn Phe Gln Lys Val Phe Asn Glu Leu Leu Val Leu Leu
635 640 645
Pro Asp Thr Ala Lys Gln Gln Glu Leu Leu Ser Ala His Thr Asp
650 655 660

Phe Cys Asn Arg Glu Lys Pro Leu Ser Thr Lys Ser Lys Lys Asn
665 670 675
Lys Lys Ser Ala Trp Gln Ala Thr Thr Gln Gln Ala Gly Leu Asp
680 685 690
Cys Arg Val Cys Pro Thr Cys Gln Gln Val Leu Ala His Gly Asp
695 700 705
Ala Ser Ser His Gln Ala Leu His Ala Ala Arg Asp Asp Asp Phe
710 715 720
Pro Ser Leu Gln Ala Ile Ala Arg Ile Ile Thr
725 730

<210> 16
<211> 299
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1748508CD1

<400> 16
Met Glu Thr Tyr Phe Val Glu Ile Ile Leu Cys Lys Tyr Val Phe
1 5 10 15
Asn Thr Tyr Phe Ile Phe Leu Thr Phe Gln Asn Tyr His Glu Ile
20 25 30
Met Thr Arg His Pro Glu Asn Tyr Gln Trp Glu Asn Trp Ser Leu
35 40 45
Glu Asn Val Ala Thr Ile Leu Ala His Arg Phe Pro Asn Ser Tyr
50 55 60
Ile Trp Val Ile Lys Cys Ser Arg Met His Leu His Lys Phe Ser
65 70 75
Cys Tyr Asp Asn Phe Val Lys Ser Asn Met Phe Gly Ala Pro Glu
80 85 90
His Asn Thr Asp Phe Gly Ala Phe Lys His Leu Tyr Met Leu Leu
95 100 105
Val Asn Ala Phe Asn Leu Ser Gln Asn Ser Leu Ser Lys Lys Ser
110 115 120
Leu Asn Val Trp Asn Lys Asp Ser Ile Ala Ser Asn Cys Arg Ser
125 130 135
Ser Pro Ser His Thr Thr Asn Gly Cys Gln Gly Glu Lys Val Arg
140 145 150
Thr Cys Glu Lys Ser Asp Glu Ser Ala Met Ser Phe Tyr Pro Pro
155 160 165
Ser Leu Asn Asp Ala Ser Phe Thr Leu Ile Gly Phe Ser Lys Gly
170 175 180
Cys Val Val Leu Asn Gln Leu Leu Phe Glu Leu Lys Glu Ala Lys
185 190 195
Lys Asp Lys Asn Ile Asp Ala Phe Ile Lys Ser Ile Arg Thr Met
200 205 210
Tyr Trp Leu Asp Gly Gly His Ser Gly Gly Ser Asn Thr Trp Val
215 220 225
Thr Tyr Pro Glu Val Leu Lys Glu Phe Ala Gln Thr Gly Ile Ile
230 235 240
Val His Thr His Val Thr Pro Tyr Gln Val Arg Asp Pro Met Arg
245 250 255
Ser Trp Ile Gly Lys Glu His Lys Lys Phe Val Gln Ile Leu Gly
260 265 270
Asp Leu Gly Met Gln Val Thr Ser Gln Ile His Phe Thr Lys Glu
275 280 285
Ala Pro Ser Ile Glu Asn His Phe Arg Val His Glu Val Phe
290 295

<210> 17

<211> 620
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2159545CD1

<400> 17
Met Ala Ser Asp Ser Met Ser Ser Lys Gln Ala Arg Asn His Ile
1   5   10   15
Thr Lys Gly Lys Arg Gln Gln Gln His Gln Gln Ile Lys Asn Arg
                20   25   30
Ser Ser Ile Ser Asp Gly Asp Gly Glu Asp Ser Phe Ile Phe Glu
                35   40   45
Ala Asn Glu Ala Trp Lys Asp Phe His Gly Ser Leu Leu Arg Phe
                50   55   60
Tyr Glu Asn Gly Glu Leu Cys Asp Val Thr Leu Lys Val Gly Ser
                65   70   75
Lys Leu Ile Ser Cys His Lys Leu Val Leu Ala Cys Val Ile Pro
                80   85   90
Tyr Phe Arg Ala Met Phe Leu Ser Glu Met Ala Glu Ala Lys Gln
                95   100   105
Thr Leu Ile Glu Ile Arg Asp Phe Asp Gly Asp Ala Ile Glu Asp
                110   115   120
Leu Val Lys Phe Val Tyr Ser Ser Arg Leu Thr Leu Thr Val Asp
                125   130   135
Asn Val Gln Pro Leu Leu Tyr Ala Ala Cys Ile Leu Gln Val Glu
                140   145   150
Leu Val Ala Arg Ala Cys Cys Glu Tyr Met Lys Leu His Phe His
                155   160   165
Pro Ser Asn Cys Leu Ala Val Arg Ala Phe Ala Glu Ser His Asn
                170   175   180
Arg Ile Asp Leu Met Asp Met Ala Asp Gln Tyr Ala Cys Asp His
                185   190   195
Phe Thr Glu Val Val Glu Cys Glu Asp Phe Val Ser Val Ser Pro
                200   205   210
Gln His Leu His Lys Leu Leu Ser Ser Ser Asp Leu Asn Ile Glu
                215   220   225
Asn Glu Lys Gln Val Tyr Asn Ala Ala Ile Lys Trp Leu Leu Ala
                230   235   240
Asn Pro Gln His His Ser Lys Trp Leu Asp Glu Thr Leu Ala Gln
                245   250   255
Val Arg Leu Pro Leu Leu Pro Val Asp Phe Leu Met Gly Val Val
                260   265   270
Ala Lys Glu Gln Ile Val Lys Gln Asn Leu Lys Cys Arg Asp Leu
                275   280   285
Leu Asp Glu Ala Arg Asn Tyr His Leu His Leu Ser Ser Arg Ala
                290   295   300
Val Pro Asp Phe Glu Tyr Ser Ile Arg Thr Thr Pro Arg Lys His
                305   310   315
Thr Ala Gly Val Leu Phe Cys Val Gly Gly Arg Gly Gly Ser Gly
                320   325   330
Asp Pro Phe Arg Ser Ile Glu Cys Tyr Ser Ile Asn Lys Asn Ser
                335   340   345
Trp Phe Phe Gly Pro Glu Met Asn Ser Arg Arg Arg His Val Gly
                350   355   360
Val Ile Ser Val Glu Gly Lys Val Tyr Ala Val Gly Gly His Asp
                365   370   375
Gly Asn Glu His Leu Gly Ser Met Glu Met Phe Asp Pro Leu Thr
                380   385   390
Asn Lys Trp Met Met Lys Ala Ser Met Asn Thr Lys Arg Arg Gly
                395   400   405

```
Ile Ala Leu Ala Ser Leu Gly Gly Pro Ile Tyr Ala Ile Gly Gly
                410                 415                 420
Leu Asp Asp Asn Thr Cys Phe Asn Asp Val Glu Arg Tyr Asp Ile
                425                 430                 435
Glu Ser Asp Gln Trp Ser Thr Val Ala Pro Met Asn Thr Pro Arg
                440                 445                 450
Gly Gly Val Gly Ser Val Ala Leu Val Asn His Val Tyr Ala Val
                455                 460                 465
Gly Gly Asn Asp Gly Met Ala Ser Leu Ser Ser Val Glu Arg Tyr
                470                 475                 480
Asp Pro His Leu Asp Lys Trp Ile Glu Val Lys Glu Met Gly Gln
                485                 490                 495
Arg Arg Ala Gly Asn Gly Val Ser Lys Leu His Gly Cys Leu Tyr
                500                 505                 510
Val Val Gly Gly Phe Asp Asp Asn Ser Pro Leu Ser Ser Val Glu
                515                 520                 525
Arg Tyr Asp Pro Arg Ser Asn Lys Trp Asp Tyr Val Ala Ala Leu
                530                 535                 540
Thr Thr Pro Arg Gly Gly Val Gly Ile Ala Thr Val Met Gly Lys
                545                 550                 555
Ile Phe Ala Val Gly Gly His Asn Gly Asn Ala Tyr Leu Asn Thr
                560                 565                 570
Val Glu Ala Phe Asp Pro Val Leu Asn Arg Trp Glu Leu Val Gly
                575                 580                 585
Ser Val Ser His Cys Arg Ala Gly Ala Gly Val Ala Val Cys Ser
                590                 595                 600
Cys Leu Thr Ser Gln Ile Arg Asp Val Gly His Gly Ser Asn Asn
                605                 610                 615
Val Val Asp Cys Met
                620
```

<210> 18
<211> 218
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8560269CD1

<400> 18

```
Met Ala Leu Val Pro Tyr Glu Glu Thr Thr Glu Phe Gly Leu Gln
 1               5                  10                  15
Lys Phe His Lys Pro Leu Ala Thr Phe Ser Phe Ala Asn His Thr
                20                  25                  30
Ile Gln Ile Arg Gln Asp Trp Arg His Leu Gly Val Ala Ala Val
                35                  40                  45
Val Trp Asp Ala Ala Ile Val Leu Ser Thr Tyr Leu Glu Met Gly
                50                  55                  60
Ala Val Glu Leu Arg Gly Arg Ser Ala Val Glu Leu Gly Ala Gly
                65                  70                  75
Thr Gly Leu Val Gly Ile Val Ala Ala Leu Leu Gly Ala His Val
                80                  85                  90
Thr Ile Thr Asp Arg Lys Val Ala Leu Glu Phe Leu Lys Ser Asn
                95                  100                 105
Val Gln Ala Asn Leu Pro Pro His Ile Gln Thr Lys Thr Val Val
                110                 115                 120
Lys Glu Leu Thr Trp Gly Gln Asn Leu Gly Ser Phe Ser Pro Gly
                125                 130                 135
Glu Phe Asp Leu Ile Leu Gly Ala Asp Ile Ile Tyr Leu Glu Glu
                140                 145                 150
Thr Phe Thr Asp Leu Leu Gln Thr Leu Glu His Leu Cys Ser Asn
                155                 160                 165
```

```
His Ser Val Ile Leu Leu Ala Cys Arg Ile Arg Tyr Glu Arg Asp
                170                 175                 180
Asn Asn Phe Leu Ala Met Leu Glu Arg Gln Phe Ile Val Arg Lys
                185                 190                 195
Val His Tyr Asp Pro Glu Lys Asp Val His Ile Tyr Glu Ala Gln
                200                 205                 210
Lys Arg Asn Gln Lys Glu Asp Leu
                215

<210> 19
<211> 427
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8710302CD1

<400> 19
Met Ala Ala Leu Ser Lys Ser Ile Pro His Asn Cys Tyr Glu Ile
 1               5                  10                  15
Gly His Thr Trp His Pro Ser Cys Arg Val Ser Phe Leu Gln Ile
                20                  25                  30
Thr Gly Gly Ala Leu Glu Glu Ser Leu Lys Ile Tyr Ala Pro Leu
                35                  40                  45
Tyr Leu Ile Ala Ala Ile Leu Arg Lys Arg Lys Leu Asp Tyr Tyr
                50                  55                  60
Leu His Lys Leu Leu Pro Glu Ile Leu Gln Ser Ala Ser Phe Leu
                65                  70                  75
Thr Ala Asn Gly Ala Leu Tyr Met Ala Phe Phe Cys Ile Leu Arg
                80                  85                  90
Arg Gly Leu Leu Thr Ile Tyr Met Ala Asn Leu Ala Thr Glu Thr
                95                 100                 105
Leu Phe Arg Met Gly Val Ala Arg Gly Thr Ile Thr Thr Leu Arg
               110                 115                 120
Asn Gly Glu Val Leu Leu Phe Cys Ile Thr Ala Ala Met Tyr Met
               125                 130                 135
Phe Phe Phe Arg Cys Lys Asp Gly Leu Lys Gly Phe Thr Phe Ser
               140                 145                 150
Ala Leu Arg Phe Ile Val Gly Lys Glu Glu Ile Pro Thr His Ser
               155                 160                 165
Phe Ser Pro Glu Ala Ala Tyr Ala Lys Val Glu Gln Lys Arg Glu
               170                 175                 180
Gln His Glu Glu Lys Pro Arg Arg Met Asn Met Ile Gly Leu Val
               185                 190                 195
Arg Lys Phe Val Asp Ser Ile Cys Lys His Gly Pro Arg His Arg
               200                 205                 210
Cys Cys Lys His Tyr Glu Asp Asn Cys Ile Ser Tyr Cys Ile Lys
               215                 220                 225
Gly Phe Ile Arg Met Phe Ser Val Gly Tyr Leu Ile Gln Cys Cys
               230                 235                 240
Leu Arg Ile Pro Ser Ala Phe Arg His Leu Phe Thr Gln Pro Ser
               245                 250                 255
Arg Leu Leu Ser Leu Phe Tyr Asn Lys Glu Asn Phe Gln Leu Gly
               260                 265                 270
Ala Phe Leu Gly Ser Phe Val Ser Ile Tyr Lys Gly Thr Ser Cys
               275                 280                 285
Phe Leu Arg Trp Ile Arg Asn Leu Asp Asp Glu Leu His Ala Ile
               290                 295                 300
Ile Ala Gly Phe Leu Ala Gly Ile Ser Met Met Phe Tyr Lys Ser
               305                 310                 315
Thr Thr Ile Ser Met Tyr Leu Ala Ser Lys Leu Val Glu Thr Met
               320                 325                 330
```

Tyr Phe Lys Gly Ile Glu Ala Gly Lys Val Pro Tyr Phe Pro His
335 340 345
Ala Asp Thr Ile Ile Tyr Ser Ile Ser Thr Ala Ile Cys Phe Gln
350 355 360
Ala Ala Val Met Glu Val Gln Thr Leu Arg Pro Ser Tyr Trp Lys
365 370 375
Phe Leu Leu Arg Leu Thr Lys Gly Lys Phe Ala Val Met Asn Arg
380 385 390
Lys Val Leu Asp Val Phe Gly Thr Gly Ala Ser Lys His Phe Gln
395 400 405
Asp Phe Ile Pro Arg Leu Asp Pro Arg Tyr Thr Thr Val Thr Pro
410 415 420
Glu Leu Pro Thr Glu Phe Ser
425

<210> 20
<211> 612
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6778214CD1

<400> 20
Met Glu Ile Ala Pro Gln Glu Ala Pro Pro Val Pro Gly Ala Asp
1 5 10 15
Gly Asp Ile Glu Glu Ala Pro Ala Glu Ala Gly Ser Pro Ser Pro
20 25 30
Ala Ser Pro Pro Ala Asp Gly Arg Leu Lys Ala Ala Ala Lys Arg
35 40 45
Val Thr Phe Pro Ser Asp Glu Asp Ile Val Ser Gly Ala Val Glu
50 55 60
Pro Lys Asp Pro Trp Arg His Ala Gln Asn Val Thr Val Asp Glu
65 70 75
Val Ile Gly Ala Tyr Lys Gln Ala Cys Gln Lys Leu Asn Cys Arg
80 85 90
Gln Ile Pro Lys Leu Leu Arg Gln Leu Gln Glu Phe Thr Asp Leu
95 100 105
Gly His Arg Leu Asp Cys Leu Asp Leu Lys Gly Glu Lys Leu Asp
110 115 120
Tyr Lys Thr Cys Glu Ala Leu Glu Glu Val Phe Lys Arg Leu Gln
125 130 135
Phe Lys Val Val Asp Leu Glu Gln Thr Asn Leu Asp Glu Asp Gly
140 145 150
Ala Ser Ala Leu Phe Asp Met Ile Glu Tyr Tyr Glu Ser Ala Thr
155 160 165
His Leu Asn Ile Ser Phe Asn Lys His Ile Gly Thr Arg Gly Trp
170 175 180
Gln Ala Ala Ala His Met Met Arg Lys Thr Ser Cys Leu Gln Tyr
185 190 195
Leu Asp Ala Arg Asn Thr Pro Leu Leu Asp His Ser Ala Pro Phe
200 205 210
Val Ala Arg Ala Leu Arg Ile Arg Ser Ser Leu Ala Val Leu His
215 220 225
Leu Glu Asn Ala Ser Leu Ser Gly Arg Pro Leu Met Leu Leu Ala
230 235 240
Thr Ala Leu Lys Met Asn Met Asn Leu Arg Glu Leu Tyr Leu Ala
245 250 255
Asp Asn Lys Leu Asn Gly Leu Gln Asp Ser Ala Gln Leu Gly Asn
260 265 270
Leu Leu Lys Phe Asn Cys Ser Leu Gln Ile Leu Asp Leu Arg Asn
275 280 285

```
Asn His Val Leu Asp Ser Gly Leu Ala Tyr Ile Cys Glu Gly Leu
                290                 295                 300
Lys Glu Gln Arg Lys Gly Leu Val Thr Leu Val Leu Trp Asn Asn
                305                 310                 315
Gln Leu Thr His Thr Gly Met Ala Phe Leu Gly Met Thr Leu Pro
                320                 325                 330
His Thr Gln Ser Leu Glu Thr Leu Asn Leu Gly His Asn Pro Ile
                335                 340                 345
Gly Asn Glu Gly Val Arg His Leu Lys Asn Gly Leu Ile Ser Asn
                350                 355                 360
Arg Ser Val Leu Arg Leu Gly Leu Ala Ser Thr Lys Leu Thr Cys
                365                 370                 375
Glu Gly Ala Val Ala Val Ala Glu Phe Ile Ala Glu Ser Pro Arg
                380                 385                 390
Leu Leu Arg Leu Asp Leu Arg Glu Asn Glu Ile Lys Thr Gly Gly
                395                 400                 405
Leu Met Ala Leu Ser Leu Ala Leu Lys Val Asn His Ser Leu Leu
                410                 415                 420
Arg Leu Asp Leu Asp Arg Glu Pro Lys Lys Glu Ala Val Lys Ser
                425                 430                 435
Phe Ile Glu Thr Gln Lys Ala Leu Leu Ala Glu Ile Gln Asn Gly
                440                 445                 450
Cys Lys Arg Asn Leu Val Leu Ala Arg Glu Arg Glu Glu Lys Glu
                455                 460                 465
Gln Pro Pro Gln Leu Ser Ala Ser Met Pro Glu Thr Thr Ala Thr
                470                 475                 480
Glu Pro Gln Pro Asp Asp Glu Pro Ala Ala Gly Val Gln Asn Gly
                485                 490                 495
Ala Pro Ser Pro Ala Pro Ser Pro Asp Ser Asp Ser Asp Ser Asp
                500                 505                 510
Ser Asp Gly Glu Glu Glu Glu Glu Glu Glu Gly Glu Arg Asp Glu
                515                 520                 525
Thr Pro Ser Gly Ala Ile Asp Thr Arg Asp Thr Gly Ser Ser Glu
                530                 535                 540
Pro Gln Pro Pro Pro Glu Pro Pro Arg Ser Gly Pro Pro Leu Pro
                545                 550                 555
Asn Gly Leu Lys Pro Glu Phe Ala Leu Ala Leu Pro Pro Glu Pro
                560                 565                 570
Pro Pro Gly Pro Glu Val Lys Gly Gly Ser Cys Gly Leu Glu His
                575                 580                 585
Glu Leu Ser Cys Ser Lys Asn Glu Lys Glu Leu Glu Glu Leu Leu
                590                 595                 600
Leu Glu Ala Ser Gln Glu Ser Gly Gln Glu Thr Leu
                605                 610

&lt;210&gt; 21
&lt;211&gt; 458
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;220&gt;
&lt;221&gt; misc_feature
&lt;223&gt; Incyte ID No: 258383CD1

&lt;400&gt; 21
Met Ala Ala Leu Ser Lys Ser Ile Pro His Asn Cys Tyr Glu Ile
  1               5                  10                  15
Gly His Thr Trp His Pro Ser Cys Arg Val Ser Phe Leu Gln Ile
                 20                  25                  30
Thr Gly Gly Ala Leu Glu Glu Ser Leu Lys Ile Tyr Ala Pro Leu
                 35                  40                  45
Tyr Leu Ile Ala Ala Ile Leu Arg Lys Arg Lys Leu Asp Tyr Tyr
                 50                  55                  60
```

Leu His Lys Leu Leu Pro Glu Ile Leu Gln Ser Ala Ser Phe Leu
65 70 75
Thr Ala Asn Gly Ala Leu Tyr Met Ala Phe Phe Cys Ile Leu Arg
80 85 90
Lys Ile Leu Gly Lys Phe Tyr Ser Trp Thr Pro Gly Phe Gly Ala
95 100 105
Ala Leu Pro Ala Ser Tyr Val Ala Ile Leu Ile Glu Arg Lys Ser
110 115 120
Arg Arg Gly Leu Leu Thr Ile Tyr Met Ala Asn Leu Ala Thr Glu
125 130 135
Thr Leu Phe Arg Met Gly Val Ala Arg Gly Thr Ile Thr Thr Leu
140 145 150
Arg Asn Gly Glu Val Leu Leu Phe Cys Ile Thr Ala Ala Met Tyr
155 160 165
Met Phe Phe Phe Arg Cys Lys Asp Gly Leu Lys Gly Phe Thr Phe
170 175 180
Ser Ala Leu Arg Phe Ile Val Gly Lys Glu Glu Ile Pro Thr His
185 190 195
Ser Phe Ser Pro Glu Ala Ala Tyr Ala Lys Val Glu Gln Lys Arg
200 205 210
Glu Gln His Glu Glu Lys Pro Arg Arg Met Asn Met Ile Gly Leu
215 220 225
Val Arg Lys Phe Val Asp Ser Ile Cys Lys His Gly Pro Arg His
230 235 240
Arg Cys Cys Lys His Tyr Glu Asp Asn Cys Ile Ser Tyr Cys Ile
245 250 255
Lys Gly Phe Ile Arg Met Phe Ser Val Gly Tyr Leu Ile Gln Cys
260 265 270
Cys Leu Arg Ile Pro Ser Ala Phe Arg His Leu Phe Thr Gln Pro
275 280 285
Ser Arg Leu Leu Ser Leu Phe Tyr Asn Lys Glu Asn Phe Gln Leu
290 295 300
Gly Ala Phe Leu Gly Ser Phe Val Ser Ile Tyr Lys Gly Thr Ser
305 310 315
Cys Phe Leu Arg Trp Ile Arg Asn Leu Asp Asp Glu Leu His Ala
320 325 330
Ile Ile Ala Gly Phe Leu Ala Gly Ile Ser Met Met Phe Tyr Lys
335 340 345
Ser Thr Thr Ile Ser Met Tyr Leu Ala Ser Lys Leu Val Glu Thr
350 355 360
Met Tyr Phe Lys Gly Ile Glu Ala Gly Lys Val Pro Tyr Phe Pro
365 370 375
His Ala Asp Thr Ile Ile Tyr Ser Ile Ser Thr Ala Ile Cys Phe
380 385 390
Gln Ala Ala Val Met Glu Val Gln Thr Leu Arg Pro Ser Tyr Trp
395 400 405
Lys Phe Leu Leu Arg Leu Thr Lys Gly Lys Phe Ala Val Met Asn
410 415 420
Arg Lys Val Leu Asp Val Phe Gly Thr Gly Ala Ser Lys His Phe
425 430 435
Gln Asp Phe Ile Pro Arg Leu Asp Pro Arg Tyr Thr Thr Val Thr
440 445 450
Pro Glu Leu Pro Thr Glu Phe Ser
455

<210> 22
<211> 1451
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2804937CD1

<400> 22

```
Met Ser Leu Val Asn Phe Glu Pro Ala Ala Arg Arg Ala Ser Asn
 1               5                  10                  15
Ile Cys Asp Thr Asp Ser His Val Ser Ser Ser Thr Ser Val Arg
                20                  25                  30
Phe Tyr Pro His Asp Val Leu Ser Leu Pro Gln Ile Arg Leu Asn
                35                  40                  45
Arg Leu Leu Thr Ile Asp Thr Asp Leu Leu Glu Gln Gln Asp Ile
                50                  55                  60
Asp Leu Ser Pro Asp Leu Ala Ala Thr Tyr Gly Pro Thr Glu Glu
                65                  70                  75
Ala Ala Gln Lys Val Lys His Tyr Tyr Arg Phe Trp Ile Leu Pro
                80                  85                  90
Gln Leu Trp Ile Gly Ile Asn Phe Asp Arg Leu Thr Leu Leu Ala
                95                  100                 105
Leu Phe Asp Arg Asn Arg Glu Ile Leu Glu Asn Val Leu Ala Val
                110                 115                 120
Ile Leu Ala Ile Leu Val Ala Phe Leu Gly Ser Ile Leu Leu Ile
                125                 130                 135
Gln Gly Phe Phe Arg Asp Ile Trp Val Phe Gln Phe Cys Leu Val
                140                 145                 150
Ile Ala Ser Cys Gln Tyr Ser Leu Leu Lys Ser Val Gln Pro Asp
                155                 160                 165
Ser Ser Ser Pro Arg His Gly His Asn Arg Ile Ile Ala Tyr Ser
                170                 175                 180
Arg Pro Val Tyr Phe Cys Ile Cys Cys Gly Leu Ile Trp Leu Leu
                185                 190                 195
Asp Tyr Gly Ser Arg Asn Leu Thr Ala Thr Lys Phe Lys Leu Tyr
                200                 205                 210
Gly Ile Thr Phe Thr Asn Pro Leu Val Phe Ile Ser Ala Arg Asp
                215                 220                 225
Leu Val Ile Val Phe Thr Leu Cys Phe Pro Ile Val Phe Phe Ile
                230                 235                 240
Gly Leu Leu Pro Gln Val Asn Thr Phe Val Met Tyr Leu Cys Glu
                245                 250                 255
Gln Leu Asp Ile His Ile Phe Gly Gly Asn Ala Thr Thr Ser Leu
                260                 265                 270
Leu Ala Ala Leu Tyr Ser Phe Ile Cys Ser Ile Val Ala Val Ala
                275                 280                 285
Leu Leu Tyr Gly Leu Cys Tyr Gly Ala Leu Lys Asp Ser Trp Asp
                290                 295                 300
Gly Gln His Ile Pro Val Leu Phe Ser Ile Phe Cys Gly Leu Leu
                305                 310                 315
Val Ala Val Ser Tyr His Leu Ser Arg Gln Ser Ser Asp Pro Ser
                320                 325                 330
Val Leu Phe Ser Leu Val Gln Ser Lys Ile Phe Pro Lys Thr Glu
                335                 340                 345
Glu Lys Asn Pro Glu Asp Pro Leu Ser Glu Val Lys Asp Pro Leu
                350                 355                 360
Pro Glu Lys Leu Arg Asn Ser Val Ser Glu Arg Leu Gln Ser Asp
                365                 370                 375
Leu Val Val Cys Ile Val Ile Gly Val Leu Tyr Phe Ala Ile His
                380                 385                 390
Val Ser Thr Val Phe Thr Val Leu Gln Pro Ala Leu Lys Tyr Val
                395                 400                 405
Leu Tyr Thr Leu Val Gly Phe Val Gly Phe Val Thr His Tyr Val
                410                 415                 420
Leu Pro Gln Val Arg Lys Gln Leu Pro Trp His Cys Phe Ser His
                425                 430                 435
Pro Leu Leu Lys Thr Leu Glu Tyr Asn Gln Tyr Glu Val Arg Asn
                440                 445                 450
Ala Ala Thr Met Met Trp Phe Glu Lys Leu His Val Trp Leu Leu
                455                 460                 465
```

Phe Val Glu Lys Asn Ile Ile Tyr Pro Leu Ile Val Leu Asn Glu
470 475 480
Leu Ser Ser Ser Ala Glu Thr Ile Ala Ser Pro Lys Lys Leu Asn
485 490 495
Thr Glu Leu Gly Ala Leu Met Ile Thr Val Ala Gly Leu Lys Leu
500 505 510
Leu Arg Ser Ser Phe Ser Ser Pro Thr Tyr Gln Tyr Val Thr Val
515 520 525
Ile Phe Thr Val Leu Phe Phe Lys Phe Asp Tyr Glu Ala Phe Ser
530 535 540
Glu Thr Met Leu Leu Asp Leu Phe Phe Met Ser Ile Leu Phe Asn
545 550 555
Lys Leu Trp Glu Leu Leu Tyr Lys Leu Gln Phe Val Tyr Thr Tyr
560 565 570
Ile Ala Pro Trp Gln Ile Thr Trp Gly Ser Ala Phe His Ala Phe
575 580 585
Ala Gln Pro Phe Ala Val Pro His Ser Ala Met Leu Phe Ile Gln
590 595 600
Ala Ala Val Ser Ala Phe Phe Ser Thr Pro Leu Asn Pro Phe Leu
605 610 615
Gly Ser Ala Ile Phe Ile Thr Ser Tyr Val Arg Pro Val Lys Phe
620 625 630
Trp Glu Arg Asp Tyr Asn Thr Lys Arg Val Asp His Ser Asn Thr
635 640 645
Arg Leu Ala Ser Gln Leu Asp Arg Asn Pro Gly Ser Asp Asp Asn
650 655 660
Asn Leu Asn Ser Ile Phe Tyr Glu His Leu Thr Arg Ser Leu Gln
665 670 675
His Ser Leu Cys Gly Asp Leu Leu Leu Gly Arg Trp Gly Asn Tyr
680 685 690
Ser Thr Gly Asp Cys Phe Ile Leu Ala Ser Asp Tyr Leu Asn Ala
695 700 705
Leu Val His Leu Ile Glu Ile Gly Asn Gly Leu Val Thr Phe Gln
710 715 720
Leu Arg Gly Leu Glu Phe Arg Gly Thr Tyr Cys Gln Gln Arg Glu
725 730 735
Val Glu Ala Ile Thr Glu Gly Val Glu Glu Asp Glu Gly Phe Cys
740 745 750
Cys Cys Glu Pro Gly His Ile Pro His Met Leu Ser Phe Asn Ala
755 760 765
Ala Phe Ser Gln Arg Trp Leu Ala Trp Glu Val Ile Val Thr Lys
770 775 780
Tyr Ile Leu Glu Gly Tyr Ser Ile Thr Asp Asn Ser Ala Ala Ser
785 790 795
Met Leu Gln Val Phe Asp Leu Arg Lys Val Leu Thr Thr Tyr Tyr
800 805 810
Val Lys Gly Ile Ile Tyr Tyr Val Thr Thr Ser Ser Lys Leu Glu
815 820 825
Glu Trp Leu Ala Asn Glu Thr Met Gln Glu Gly Leu Arg Leu Cys
830 835 840
Ala Asp Arg Asn Tyr Val Asp Val Asp Pro Thr Phe Asn Pro Asn
845 850 855
Ile Asp Glu Asp Tyr Asp His Arg Leu Ala Gly Ile Ser Arg Glu
860 865 870
Ser Phe Cys Val Ile Tyr Leu Asn Trp Ile Glu Tyr Cys Ser Ser
875 880 885
Arg Arg Ala Lys Pro Val Asp Val Asp Lys Asp Ser Ser Leu Val
890 895 900
Thr Leu Cys Tyr Gly Leu Cys Val Leu Gly Arg Arg Ala Leu Gly
905 910 915
Thr Ala Ser His His Met Ser Ser Asn Leu Glu Ser Phe Leu Tyr
920 925 930
Gly Leu His Ala Leu Phe Lys Gly Asp Phe Arg Ile Ser Ser Ile

935 940 945
Arg Asp Glu Trp Ile Phe Ala Asp Met Glu Leu Leu Arg Lys Val
950 955 960
Val Val Pro Gly Ile Arg Met Ser Ile Lys Leu His Gln Asp His
965 970 975
Phe Thr Ser Pro Asp Glu Tyr Asp Asp Pro Thr Val Leu Tyr Glu
980 985 990
Ala Ile Val Ser His Glu Lys Asn Leu Val Ile Ala His Glu Gly
995 1000 1005
Asp Pro Ala Trp Arg Ser Ala Val Leu Ala Asn Ser Pro Ser Leu
1010 1015 1020
Leu Ala Leu Arg His Val Met Asp Asp Gly Thr Asn Glu Tyr Lys
1025 1030 1035
Ile Ile Met Leu Asn Arg Arg Tyr Leu Ser Phe Arg Val Ile Lys
1040 1045 1050
Val Asn Lys Glu Cys Val Arg Gly Leu Trp Ala Gly Gln Gln Gln
1055 1060 1065
Glu Leu Val Phe Leu Arg Asn Arg Asn Pro Glu Arg Gly Ser Ile
1070 1075 1080
Gln Asn Ala Lys Gln Ala Leu Arg Asn Met Ile Asn Ser Ser Cys
1085 1090 1095
Asp Gln Pro Ile Gly Tyr Pro Ile Phe Val Ser Pro Leu Thr Thr
1100 1105 1110
Ser Tyr Ser Asp Ser His Glu Gln Leu Lys Asp Ile Leu Gly Gly
1115 1120 1125
Pro Ile Ser Leu Gly Asn Ile Arg Asn Phe Ile Val Ser Thr Trp
1130 1135 1140
His Arg Leu Arg Lys Gly Cys Gly Ala Gly Cys Asn Ser Gly Gly
1145 1150 1155
Asn Ile Glu Asp Ser Asp Thr Gly Gly Gly Thr Ser Cys Thr Gly
1160 1165 1170
Asn Asn Ala Thr Thr Ala Asn Asn Pro His Ser Asn Val Thr Gln
1175 1180 1185
Gly Ser Ile Gly Asn Pro Gly Gln Gly Ser Gly Thr Gly Leu His
1190 1195 1200
Pro Pro Val Thr Ser Tyr Pro Pro Thr Leu Gly Thr Ser His Ser
1205 1210 1215
Ser His Ser Val Gln Ser Gly Leu Val Arg Gln Ser Pro Ala Arg
1220 1225 1230
Ala Ser Val Ala Ser Gln Ser Ser Tyr Cys Tyr Ser Ser Arg His
1235 1240 1245
Ser Ser Leu Arg Met Ser Thr Thr Gly Phe Val Pro Cys Arg Arg
1250 1255 1260
Ser Ser Thr Ser Gln Ile Ser Leu Arg Asn Leu Pro Ser Ser Ile
1265 1270 1275
Gln Ser Arg Leu Ser Met Val Asn Gln Met Glu Pro Ser Gly Gln
1280 1285 1290
Ser Gly Leu Ala Cys Val Gln His Gly Leu Pro Ser Ser Ser Ser
1295 1300 1305
Ser Ser Gln Ser Ile Pro Ala Cys Lys His His Thr Leu Val Gly
1310 1315 1320
Phe Leu Ala Thr Glu Gly Gly Gln Ser Ser Ala Thr Asp Ala Gln
1325 1330 1335
Pro Gly Asn Thr Leu Ser Pro Ala Asn Asn Ser His Ser Arg Lys
1340 1345 1350
Ala Glu Val Ile Tyr Arg Val Gln Ile Val Asp Pro Ser Gln Ile
1355 1360 1365
Leu Glu Gly Ile Asn Leu Ser Lys Arg Lys Glu Leu Gln Trp Pro
1370 1375 1380
Asp Glu Gly Ile Arg Leu Lys Ala Gly Arg Asn Ser Trp Lys Asp
1385 1390 1395
Trp Ser Pro Gln Glu Gly Met Glu Gly His Val Ile His Arg Trp
1400 1405 1410

Val Pro Cys Ser Arg Asp Pro Gly Thr Arg Ser His Ile Asp Lys
1415 1420 1425
Ala Val Leu Leu Val Gln Ile Asp Asp Lys Tyr Val Thr Val Ile
1430 1435 1440
Glu Thr Gly Val Leu Glu Leu Gly Ala Glu Val
1445 1450

<210> 23
<211> 184
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7494915CD1

<400> 23
Met Met Pro Gly Glu Lys Lys Gln Ser Gly Ser Gln Gln Gln Asn
1 5 10 15
Asn Gly Gln Ala Ser Lys Asn Arg Thr Gln Lys Glu Ile Val Thr
20 25 30
Gln Lys Arg Pro Ile Thr Ser Asn Glu Ile Glu Leu Val Val Lys
35 40 45
Lys Lys Leu Pro Arg Glu Lys Gly Pro Gly Pro Asp Gly Phe Ile
50 55 60
Ala Glu Phe Phe Arg Thr Val Lys Glu Glu Leu Glu Pro Thr Leu
65 70 75
Leu Lys Leu Phe Gln Lys Ile Glu Arg Glu Arg Ile Leu Pro Asn
80 85 90
Thr Phe Tyr Gly Val Ser Ile Thr Leu Met Pro Lys Pro Glu Lys
95 100 105
Asp Thr Thr Ala Thr Thr Thr Thr Thr Thr Thr Asn Tyr Arg Pro
110 115 120
Thr Ser Leu Met Asn Val Asp Ser Lys Ile Leu Asn Lys Ile Leu
125 130 135
Ala Asn Gln Ile Gln Pro His Ile Lys Lys Ile Ile His His Asn
140 145 150
Gln Lys Leu Phe Ser Leu Ile Arg Ser His Leu Ser Ile Leu Ala
155 160 165
Phe Val Ala Ile Ala Phe Gly Val Leu Asp Val Lys Pro Leu Pro
170 175 180
Ile Pro Met His

<210> 24
<211> 407
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2073751CD1

<400> 24
Met Ala Ala Glu Ile Asp Phe Leu Arg Glu Gln Asn Arg Arg Leu
1 5 10 15
Asn Glu Asp Phe Arg Arg Tyr Gln Met Glu Ser Phe Ser Lys Tyr
20 25 30
Ser Ser Val Gln Lys Ala Val Cys Gln Gly Glu Gly Asp Asp Thr
35 40 45
Phe Glu Asn Leu Val Phe Asp Gln Ser Phe Leu Ala Pro Leu Val
50 55 60
Thr Glu Tyr Asp Lys His Leu Gly Glu Leu Asn Gly Gln Leu Lys

```
                65                  70                  75
Tyr Tyr Gln Lys Gln Val Gly Glu Met Lys Leu Gln Phe Glu Asn
                80                  85                  90
Val Ile Lys Glu Asn Glu Arg Leu His Ser Glu Leu Lys Asp Ala
                95                  100                 105
Val Glu Lys Lys Leu Glu Ala Phe Pro Leu Gly Thr Glu Val Gly
                110                 115                 120
Thr Asp Ile Tyr Ala Asp Asp Glu Thr Val Arg Asn Leu Gln Glu
                125                 130                 135
Gln Leu Gln Leu Ala Asn Gln Glu Lys Thr Gln Ala Val Glu Leu
                140                 145                 150
Trp Gln Thr Val Ser Gln Glu Leu Asp Arg Leu His Lys Leu Tyr
                155                 160                 165
Gln Glu His Met Thr Glu Ala Gln Ile His Val Phe Glu Ser Gln
                170                 175                 180
Lys Gln Lys Asp Gln Leu Phe Asp Phe Gln Gln Leu Thr Lys Gln
                185                 190                 195
Leu His Val Thr Asn Glu Asn Met Glu Val Thr Asn Gln Gln Phe
                200                 205                 210
Leu Lys Thr Val Thr Glu Gln Ser Val Ile Ile Glu Gln Leu Arg
                215                 220                 225
Lys Lys Leu Arg Gln Ala Lys Leu Glu Leu Arg Val Ala Val Ala
                230                 235                 240
Lys Val Glu Glu Leu Thr Asn Val Thr Glu Asp Leu Gln Gly Gln
                245                 250                 255
Met Lys Lys Lys Glu Lys Asp Val Val Ser Ala His Gly Arg Glu
                260                 265                 270
Glu Ala Ser Asp Arg Arg Leu Gln Gln Leu Gln Ser Ser Ile Lys
                275                 280                 285
Gln Leu Glu Ile Arg Leu Cys Val Thr Ile Gln Glu Ala Asn Gln
                290                 295                 300
Leu Arg Thr Glu Asn Thr His Leu Glu Lys Gln Thr Arg Glu Leu
                305                 310                 315
Gln Ala Lys Cys Asn Glu Leu Glu Asn Glu Arg Tyr Glu Ala Ile
                320                 325                 330
Val Arg Ala Arg Asn Ser Met Gln Leu Leu Glu Glu Ala Asn Leu
                335                 340                 345
Gln Lys Ser Gln Ala Leu Leu Glu Glu Lys Gln Lys Glu Glu Asp
                350                 355                 360
Ile Glu Lys Met Lys Glu Thr Val Ser Arg Phe Val Gln Asp Ala
                365                 370                 375
Thr Ile Arg Thr Lys Lys Glu Val Ala Asn Thr Lys Lys Gln Cys
                380                 385                 390
Asn Ile Gln Ile Ser Arg Leu Thr Glu Glu Leu Ser Ala Leu Gln
                395                 400                 405
Met Glu
```

<210> 25
<211> 261
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3178841CD1

<400> 25

```
Met Thr Cys Leu Ala Pro Thr Met Ser Ala Glu Leu Asn Val Pro
 1               5                  10                  15
Ile Asp Pro Ser Ala Pro Ala Cys Pro Glu Pro Gly His Lys Gly
                20                  25                  30
Met Asp Tyr Arg Asp Trp Val Arg Arg Ser Tyr Leu Glu Leu Val
```

35 40 45
Thr Ser Asn His His Ser Val Gln Ala Leu Ser Trp Arg Lys Leu
50 55 60
Tyr Leu Ser Arg Ala Lys Leu Lys Ala Ser Ser Arg Thr Ser Ala
65 70 75
Leu Leu Ser Gly Phe Ala Met Val Ala Met Val Glu Val Gln Leu
80 85 90
Glu Thr Gln Tyr Gln Tyr Pro Arg Pro Leu Leu Ile Ala Phe Ser
95 100 105
Ala Cys Thr Thr Val Leu Val Ala Val His Leu Phe Ala Leu Leu
110 115 120
Ile Ser Thr Cys Ile Leu Pro Asn Val Glu Ala Val Ser Asn Ile
125 130 135
His Asn Leu Asn Ser Ile Ser Glu Ser Pro His Glu Arg Met His
140 145 150
Pro Tyr Ile Glu Leu Ala Trp Gly Phe Ser Thr Val Leu Gly Ile
155 160 165
Leu Leu Phe Leu Ala Glu Val Val Leu Leu Cys Trp Ile Lys Phe
170 175 180
Leu Pro Val Asp Ala Arg Arg Gln Pro Gly Pro Pro Pro Gly Pro
185 190 195
Gly Ser His Thr Gly Trp Gln Ala Ala Leu Val Ser Thr Ile Ile
200 205 210
Met Val Pro Val Gly Leu Ile Phe Val Val Phe Thr Ile His Phe
215 220 225
Tyr Arg Ser Leu Val Arg His Lys Thr Glu Arg His Asn Arg Glu
230 235 240
Ile Glu Glu Leu His Lys Leu Lys Val Gln Leu Asp Gly His Glu
245 250 255
Arg Ser Leu Gln Val Leu
260

<210> 26
<211> 209
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3674807CD1

<400> 26
Met Ala Thr Ile Ala Ala Ala Ala Phe Glu Ala Leu Met Asp Gly
1 5 10 15
Val Thr Cys Trp Asp Val Pro Arg Gly Pro Ile Pro Ser Glu Leu
20 25 30
Leu Leu Ile Gly Glu Ala Ala Phe Pro Val Met Val Asn Asp Lys
35 40 45
Gly Gln Val Leu Ile Ala Ala Ser Ser Tyr Gly Arg Gly Arg Leu
50 55 60
Val Val Val Ser His Glu Gly Tyr Leu Ser His Ala Gly Leu Ala
65 70 75
Pro Phe Leu Leu Asn Ala Val Ser Trp Leu Cys Pro Cys Pro Gly
80 85 90
Ala Pro Val Gly Val His Pro Ser Leu Ala Pro Leu Val Asn Ile
95 100 105
Leu Gln Asp Ala Gly Leu Glu Ala Gln Val Lys Pro Glu Pro Gly
110 115 120
Glu Pro Leu Gly Val Tyr Cys Ile Asn Ala Tyr Asn Asp Thr Leu
125 130 135
Thr Ala Thr Leu Ile Gln Phe Val Lys His Gly Gly Gly Leu Leu
140 145 150
Ile Gly Gly Gln Ala Trp Tyr Trp Ala Ser Gln His Gly Pro Asp

```
                  155                 160                 165
Lys Val Leu Ser Arg Phe Pro Gly Asn Lys Val Thr Ser Val Ala
                  170                 175                 180
Gly Val Tyr Phe Thr Asp Thr Tyr Gly Asp Arg Asp Arg Phe Lys
                  185                 190                 195
Val Ser Lys Lys Val Pro Lys Ile Pro Leu His Val Arg Arg
                  200                 205
```

<210> 27
<211> 333
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1794922CD1

<400> 27

```
Met Lys Met Asp Val Ser Val Arg Ala Ala Gly Cys Ser Asp Asp
  1               5                  10                  15
Leu Ser Ser Gly Glu Ala Asp Val Asp Pro Lys Leu Leu Glu Leu
                 20                  25                  30
Thr Ala Asp Glu Glu Lys Cys Arg Ser Ile Arg Arg Gln Tyr Arg
                 35                  40                  45
Gln Leu Met Tyr Cys Val Arg Gln Asn Arg Glu Asp Ile Val Ser
                 50                  55                  60
Ser Ala Asn Asn Ser Leu Thr Glu Ala Leu Glu Glu Ala Asn Val
                 65                  70                  75
Leu Phe Asp Gly Val Ser Arg Thr Arg Glu Ala Ala Leu Asp Ala
                 80                  85                  90
Arg Phe Leu Val Met Ala Ser Asp Leu Gly Lys Glu Lys Ala Lys
                 95                 100                 105
Gln Leu Asn Ser Asp Met Asn Phe Phe Asn Gln Leu Ala Phe Cys
                110                 115                 120
Asp Phe Leu Phe Leu Phe Val Gly Leu Asn Trp Met Glu Gly Asp
                125                 130                 135
Pro Asp Lys Leu Ser Asp Cys Asp Asp Ser Ile Ala Leu Ser Phe
                140                 145                 150
Trp Lys Ala Ile Glu Lys Glu Ala Thr Ser Trp Met Val Lys Ala
                155                 160                 165
Glu Thr Phe His Phe Val Phe Gly Ser Phe Lys Leu Glu Arg Ser
                170                 175                 180
Ala Pro Lys Pro Arg Leu Glu His Gln Lys Lys Val Arg Lys Met
                185                 190                 195
Glu Glu Asn Gly Asn Met Pro Thr Lys Leu Gln Lys Leu Asp Leu
                200                 205                 210
Ser Ser Tyr Pro Glu Ala Thr Glu Lys Asn Val Glu Arg Ile Leu
                215                 220                 225
Gly Leu Leu Gln Thr Tyr Phe Arg Lys Tyr Pro Asp Thr Pro Val
                230                 235                 240
Ser Tyr Phe Glu Phe Val Ile Asp Pro Asn Ser Phe Ser Arg Thr
                245                 250                 255
Val Glu Asn Ile Phe Tyr Val Ser Phe Ile Val Arg Asp Gly Phe
                260                 265                 270
Ala Arg Ile Arg Leu Asp Glu Asp Arg Leu Pro Ile Leu Glu Pro
                275                 280                 285
Met Asn Val Asn Gln Met Gly Glu Gly Asn Asp Ser Ser Cys His
                290                 295                 300
Gly Arg Lys Gln Gly Val Ile Ser Leu Thr Leu Gln Glu Trp Lys
                305                 310                 315
Asn Ile Val Ala Ala Phe Glu Ile Ser Glu Ala Met Ile Thr Tyr
                320                 325                 330
Ser Ser Tyr
```

<210> 28
<211> 257
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1795509CD1

<400> 28
Met Val Ala Glu Lys Glu Thr Leu Ser Leu Asn Lys Cys Pro Asp
1 5 10 15
Lys Met Pro Lys Arg Thr Lys Leu Leu Ala Gln Gln Pro Leu Pro
20 25 30
Val His Gln Pro His Ser Leu Val Ser Glu Gly Phe Thr Val Lys
35 40 45
Ala Met Met Lys Asn Ser Val Val Arg Gly Pro Pro Ala Ala Gly
50 55 60
Ala Phe Lys Glu Arg Pro Thr Lys Pro Thr Ala Phe Arg Lys Phe
65 70 75
Tyr Glu Arg Gly Asp Phe Pro Ile Ala Leu Glu His Asp Ser Lys
80 85 90
Gly Asn Lys Ile Ala Trp Lys Val Glu Ile Glu Lys Leu Asp Tyr
95 100 105
His His Tyr Leu Pro Leu Phe Phe Asp Gly Leu Cys Glu Met Thr
110 115 120
Phe Pro Tyr Glu Phe Phe Ala Arg Gln Gly Ile His Asp Met Leu
125 130 135
Glu His Gly Gly Asn Lys Ile Leu Pro Val Leu Pro Gln Leu Ile
140 145 150
Ile Pro Ile Lys Asn Ala Leu Asn Leu Arg Asn Arg Gln Val Ile
155 160 165
Cys Val Thr Leu Lys Val Leu Gln His Leu Val Val Ser Ala Glu
170 175 180
Met Val Gly Lys Ala Leu Val Pro Tyr Tyr Arg Gln Ile Leu Pro
185 190 195
Val Leu Asn Ile Phe Lys Asn Met Asn Val Asn Ser Gly Asp Gly
200 205 210
Ile Asp Tyr Ser Gln Gln Lys Arg Glu Asn Ile Gly Asp Leu Ile
215 220 225
Gln Glu Thr Leu Glu Ala Phe Glu Arg Tyr Gly Gly Glu Asn Ala
230 235 240
Phe Ile Asn Ile Lys Tyr Val Val Pro Thr Tyr Glu Ser Cys Leu
245 250 255
Leu Asn

<210> 29
<211> 293
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2017180CD1

<400> 29
Met Arg Val Asp Ser Ser Ala Asp Pro Thr Met Ser Gln Glu Gln
1 5 10 15
Gly Pro Gly Ser Ser Thr Pro Pro Ser Ser Pro Thr Leu Leu Asp
20 25 30

```
Ala Leu Leu Gln Asn Leu Tyr Asp Phe Gly Gly Thr Glu Gly Glu
                35                  40                  45
Thr Glu Gln Lys Lys Ile Ile Lys Lys Arg Glu Asn Lys Lys Arg
                50                  55                  60
Asp Val Met Ala Ser Ala Ala Leu Ala Ala Glu Pro Ser Pro Leu
                65                  70                  75
Pro Gly Ser Leu Ile Arg Gly Gln Arg Lys Ser Ala Ser Ser Phe
                80                  85                  90
Phe Lys Glu Leu Arg Glu Glu Arg His Cys Ala Pro Ser Gly Thr
                95                  100                 105
Pro Thr Gly Pro Glu Ile Leu Ala Ala Ala Val Pro Pro Ser Ser
                110                 115                 120
Leu Lys Asn Asn Arg Glu Gln Val Glu Val Val Glu Phe His Ser
                125                 130                 135
Asn Lys Lys Arg Lys Leu Thr Pro Asp His Asn Lys Asn Thr Lys
                140                 145                 150
Gln Ala Asn Pro Ser Val Leu Glu Arg Asp Val Asp Thr Gln Glu
                155                 160                 165
Phe Asn Leu Glu Lys Ala Arg Leu Glu Val His Arg Phe Gly Ile
                170                 175                 180
Thr Gly Tyr Gly Lys Gly Lys Glu Arg Ile Leu Glu Gln Glu Arg
                185                 190                 195
Ala Ile Met Leu Gly Ala Lys Pro Pro Lys Lys Ser Tyr Val Asn
                200                 205                 210
Tyr Lys Val Leu Gln Glu Gln Ile Lys Glu Lys Lys Ala Ala Lys
                215                 220                 225
Glu Glu Glu Lys Arg Leu Ala Gln Glu Thr Asp Ile Phe Lys Lys
                230                 235                 240
Lys Lys Arg Lys Gly Gln Glu Asp Arg Lys Ser Lys Lys Lys Ser
                245                 250                 255
Ala Pro Ser Ile Leu Ser Asn Gly Arg Ile Gly Gln Val Gly Lys
                260                 265                 270
Phe Lys Asn Gly Thr Leu Ile Leu Ser Pro Val Asp Ile Lys Lys
                275                 280                 285
Ile Asn Ser Ser Arg Val Ala Lys
                290
```

<210> 30
<211> 598
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 219442CD1

<400> 30

```
Met Ala Ala Ser Val Thr Asn Ala Val Pro Pro His Asn Phe Lys
  1               5                   10                  15
Ser Gln Glu Val Thr Pro Ala Cys Leu Asp Gly Lys Ser Leu Arg
                20                  25                  30
Ala Gly Ile Thr Glu Val Lys Glu Pro Ser Val Thr Ser Pro Thr
                35                  40                  45
Pro Ser Asp Ile Gln Gln Asn Lys Gly Leu Pro Lys Pro Glu Phe
                50                  55                  60
Arg Phe Lys Gly Gln Ser Thr Lys Ser Asp Ser Ala Glu Asp Tyr
                65                  70                  75
Leu Leu Trp Lys Arg Leu Gln Gly Val Ser Ala Ala Cys Pro Ala
                80                  85                  90
Pro Ser Ser Ala Ala His Gln Leu Glu His Leu Ser Ala Lys Leu
                95                  100                 105
Gln Lys Ile Asp Glu Gln Leu Leu Ala Ile Gln Asn Ile Ala Glu
                110                 115                 120
```

Asn Ile Glu Gln Asp Phe Pro Lys Pro Glu Met Leu Asp Leu His
125 130 135
Cys Asp Lys Ile Gly Pro Val Asp His Ile Glu Phe Ser Ser Gly
140 145 150
Pro Glu Phe Lys Lys Thr Leu Ala Ser Lys Thr Ile Ser Ile Ser
155 160 165
Glu Glu Val Arg Phe Leu Thr His Met Asp Glu Glu Asp Gln Ser
170 175 180
Asp Lys Lys Glu Thr Ser Glu Pro Glu Phe Ser Ile Thr Glu Asn
185 190 195
Tyr Ser Gly Gln Lys Thr Cys Val Phe Pro Thr Ala Asp Ser Ala
200 205 210
Val Ser Leu Ser Ser Ser Ser Asp Gln Asn Thr Thr Ser Pro Gly
215 220 225
Met Asn Ser Ser Asp Glu Leu Cys Glu Ser Val Ser Val His Pro
230 235 240
Leu Gln Met Thr Gly Leu Thr Asp Ile Ala Asp Ile Ile Asp Asp
245 250 255
Leu Ile Ile Lys Asp Gly Val Ser Ser Glu Glu Leu Gly Leu Thr
260 265 270
Glu Gln Ala Met Gly Thr Ser Arg Ile Gln His Tyr Ser Gly Arg
275 280 285
His Ser Gln Arg Thr Asp Lys Glu Arg Arg Glu Ile Gln Ala Trp
290 295 300
Met Lys Arg Lys Arg Lys Glu Arg Met Ala Lys Tyr Leu Asn Glu
305 310 315
Leu Ala Glu Lys Arg Gly Gln Glu His Asp Pro Phe Cys Pro Arg
320 325 330
Ser Asn Pro Leu Tyr Met Thr Ser Arg Glu Ile Arg Leu Arg Gln
335 340 345
Lys Met Lys His Glu Lys Asp Arg Leu Leu Leu Ser Glu His Tyr
350 355 360
Ser Arg Arg Ile Ser Gln Ala Tyr Gly Leu Met Asn Glu Leu Leu
365 370 375
Ser Glu Ser Val Gln Leu Pro Thr Leu Pro Gln Lys Pro Leu Pro
380 385 390
Asn Lys Pro Ser Pro Thr Gln Ser Ser Ser Cys Gln His Cys Pro
395 400 405
Ser Pro Arg Gly Glu Asn Gln His Gly His Ser Phe Leu Ile Asn
410 415 420
Arg Pro Gly Lys Val Lys Tyr Met Ser Lys Pro Ser Tyr Ile His
425 430 435
Lys Arg Lys Ser Phe Gly Gln Pro Gln Gly Ser Pro Trp Pro His
440 445 450
Gly Thr Ala Thr Phe Thr Ile Gln Lys Lys Ala Gly Gly Ala Lys
455 460 465
Ala Ala Val Arg Lys Ala Thr Gln Ser Pro Val Thr Phe Gln Lys
470 475 480
Gly Ser Asn Ala Pro Cys His Ser Leu Gln His Thr Lys Lys His
485 490 495
Gly Ser Ala Gly Leu Ala Pro Gln Thr Lys Gln Val Cys Val Glu
500 505 510
Tyr Glu Arg Glu Glu Thr Val Val Ser Pro Trp Thr Ile Pro Ser
515 520 525
Glu Ile His Lys Ile Leu His Glu Ser His Asn Ser Leu Leu Gln
530 535 540
Asp Leu Ser Pro Thr Glu Glu Glu Glu Pro Glu His Pro Phe Gly
545 550 555
Val Gly Gly Val Asp Ser Val Ser Glu Ser Thr Gly Ser Ile Leu
560 565 570
Ser Lys Leu Asp Trp Asn Ala Ile Glu Asp Met Val Ala Ser Val
575 580 585
Glu Asp Gln Gly Leu Ser Val His Trp Ala Leu Asp Leu

590 595

<210> 31
<211> 470
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2597459CD1

<400> 31
Met Pro Ser Glu Arg Cys Leu Ser Ile Gln Glu Met Leu Thr Gly
1 5 10 15
Gln Arg Leu Cys His Ser Glu Ser His Asn Asp Ser Val Leu Ala
20 25 30
Ala Leu Asn Gln Gln Arg Ser Asp Gly Ile Leu Cys Asp Ile Thr
35 40 45
Leu Ile Ala Glu Glu Gln Lys Phe His Ala His Lys Ala Val Leu
50 55 60
Ala Ala Cys Ser Asp Tyr Phe Arg Ala Met Phe Ser Leu Cys Met
65 70 75
Val Glu Ser Gly Ala Asp Glu Val Asn Leu His Gly Val Thr Ser
80 85 90
Leu Gly Leu Lys Gln Ala Leu Glu Phe Ala Tyr Thr Gly Gln Ile
95 100 105
Leu Leu Glu Pro Gly Val Ile Gln Asp Val Leu Ala Ala Gly Ser
110 115 120
His Leu Gln Leu Leu Glu Leu Leu Asn Leu Cys Ser His Tyr Leu
125 130 135
Ile Gln Glu Leu Asn Ser Phe Asn Tyr Leu Asp Leu Tyr Arg Leu
140 145 150
Ala Asp Leu Phe Asn Leu Thr Leu Leu Glu Lys Ala Val Ile Asp
155 160 165
Phe Leu Val Lys His Leu Ser Glu Leu Leu Lys Ser Arg Pro Glu
170 175 180
Glu Val Leu Thr Leu Pro Tyr Cys Leu Leu Gln Glu Val Leu Lys
185 190 195
Ser Asp Arg Leu Thr Ser Leu Ser Glu Glu Gln Ile Trp Gln Leu
200 205 210
Ala Val Arg Trp Leu Glu His Asn Cys His Tyr Gln Tyr Met Asp
215 220 225
Glu Leu Leu Gln Tyr Ile Arg Phe Gly Leu Met Asp Val Asp Thr
230 235 240
Leu His Thr Val Ala Leu Ser His Pro Leu Val Gln Ala Ser Glu
245 250 255
Thr Ala Thr Ala Leu Val Asn Glu Ala Leu Glu Tyr His Gln Ser
260 265 270
Ile Tyr Ala Gln Pro Val Trp Gln Thr Arg Arg Thr Lys Pro Arg
275 280 285
Phe Gln Ser Asp Thr Leu Tyr Ile Ile Gly Gly Lys Lys Arg Glu
290 295 300
Val Cys Lys Val Lys Glu Leu Arg Tyr Phe Asn Pro Val Asp Gln
305 310 315
Glu Asn Ala Leu Ile Ala Ala Ile Ala Asn Trp Ser Glu Leu Ala
320 325 330
Pro Met Pro Val Gly Arg Ser His His Cys Val Ala Val Met Gly
335 340 345
Asp Phe Leu Phe Val Ala Gly Gly Glu Val Glu His Ala Ser Gly
350 355 360
Arg Thr Cys Ala Val Arg Thr Ala Cys Arg Tyr Asp Pro Arg Ser
365 370 375
Asn Ser Trp Ala Glu Ile Ala Pro Met Lys Asn Cys Arg Glu His

```
                380                     385                     390
Phe Val Leu Gly Ala Met Glu Glu Tyr Leu Tyr Ala Val Gly Gly
                395                     400                     405
Arg Asn Glu Leu Arg Gln Val Leu Pro Thr Val Glu Arg Tyr Cys
                410                     415                     420
Pro Lys Lys Asn Lys Trp Thr Phe Val Gln Ser Phe Asp Arg Ser
                425                     430                     435
Leu Ser Cys His Ala Gly Tyr Val Ala Asp Gly Leu Leu Trp Ile
                440                     445                     450
Ser Gly Arg Thr Tyr Leu Met Leu Asp Leu Ser Lys His Thr Phe
                455                     460                     465
Ile Val Val Tyr Ile
                470
```

<210> 32
<211> 311
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2783863CD1

<400> 32

```
Met His Gln Lys Leu Leu Lys Ser Ala His Tyr Ile Glu Leu Gly
 1               5                  10                  15
Ser Tyr Gln Tyr Trp Pro Val Leu Val Pro Arg Gly Ile Arg Leu
                20                  25                  30
Tyr Thr Tyr Glu Gln Ile Pro Gly Ser Leu Lys Asp Asn Pro Tyr
                35                  40                  45
Ile Thr Asp Gly Tyr Arg Ala Tyr Leu Pro Ser Arg Leu Cys Ile
                50                  55                  60
Lys Ser Leu Phe Ile Leu Ser Asn Glu Thr Val Asn Ile Trp Ser
                65                  70                  75
His Leu Leu Gly Phe Phe Leu Phe Phe Thr Leu Gly Ile Tyr Asp
                80                  85                  90
Met Thr Ser Val Leu Pro Ser Ala Ser Ala Ser Arg Glu Asp Phe
                95                 100                 105
Val Ile Cys Ser Ile Cys Leu Phe Cys Phe Gln Val Cys Met Leu
               110                 115                 120
Cys Ser Val Gly Tyr His Leu Phe Ser Cys His Arg Ser Glu Lys
               125                 130                 135
Thr Cys Arg Arg Trp Met Ala Leu Asp Tyr Ala Gly Ile Ser Ile
               140                 145                 150
Gly Ile Leu Gly Cys Tyr Val Ser Gly Val Phe Tyr Ala Phe Tyr
               155                 160                 165
Cys Asn Asn Tyr Trp Arg Gln Val Tyr Leu Ile Thr Val Leu Ala
               170                 175                 180
Met Ile Leu Ala Val Phe Phe Ala Gln Ile His Pro Asn Tyr Leu
               185                 190                 195
Thr Gln Gln Trp Gln Arg Leu Arg Ser Ile Ile Phe Cys Ser Val
               200                 205                 210
Ser Gly Tyr Gly Val Ile Pro Thr Leu His Trp Val Trp Leu Asn
               215                 220                 225
Gly Gly Ile Gly Ala Pro Ile Val Gln Asp Phe Ala Pro Arg Val
               230                 235                 240
Ile Val Met Tyr Met Ile Ala Leu Leu Ala Phe Leu Phe Tyr Ile
               245                 250                 255
Ser Lys Val Pro Glu Arg Tyr Phe Pro Gly Gln Leu Asn Tyr Leu
               260                 265                 270
Gly Ser Ser His Gln Ile Trp His Ile Leu Ala Val Val Met Leu
               275                 280                 285
Tyr Trp Trp His Gln Ser Thr Val Tyr Val Met Gln Tyr Arg His
```

290 295 300
Ser Lys Pro Cys Pro Asp Tyr Val Ser His Leu
305 310

<210> 33
<211> 894
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2902971CD1

<400> 33
Met Ala Thr Ser Met Ala Ala Ala Ser Gly Arg Phe Glu Ser Ala
1 5 10 15
Lys Ser Ile Glu Glu Arg Lys Glu Gln Thr Arg Asn Ala Arg Ala
20 25 30
Glu Val Leu Arg Gln Ala Lys Ala Asn Phe Glu Lys Glu Glu Arg
35 40 45
Arg Lys Glu Leu Lys Arg Leu Arg Gly Glu Asp Thr Trp Met Leu
50 55 60
Pro Asp Val Asn Glu Arg Ile Glu Gln Phe Ser Gln Glu His Ser
65 70 75
Val Lys Lys Lys Lys Lys Lys Asp Lys His Ser Lys Lys Ala Lys
80 85 90
Lys Glu Lys Lys Lys Lys Ser Lys Lys Gln Lys Tyr Glu Lys Asn
95 100 105
Asn Glu Ser Ser Asp Ser Ser Ser Ser Ser Glu Asp Glu Trp Val
110 115 120
Glu Ala Val Pro Ser Gln Thr Pro Asp Lys Glu Lys Ala Trp Lys
125 130 135
Val Lys Asp Glu Lys Ser Gly Lys Asp Asp Thr Gln Ile Ile Lys
140 145 150
Arg Asp Glu Trp Met Thr Val Asp Phe Met Ser Val Lys Thr Val
155 160 165
Ser Ser Ser Ser Leu Lys Ala Glu Lys Glu Thr Met Arg Lys Ile
170 175 180
Glu Gln Glu Lys Asn Gln Ala Leu Glu Gln Ser Lys Leu Met Glu
185 190 195
Arg Glu Leu Asn Pro Tyr Trp Lys Asp Gly Gly Thr Gly Leu Pro
200 205 210
Pro Glu Asp Cys Ser Val Ser Ser Ile Thr Lys Val Ser Val Val
215 220 225
Glu Asp Gly Gly Leu Ser Trp Leu Arg Lys Ser Tyr Leu Arg Met
230 235 240
Lys Glu Gln Ala Glu Lys Gln Ser Arg Asn Phe Glu Asp Ile Val
245 250 255
Ala Glu Arg Tyr Gly Ser Met Glu Ile Phe Gln Ser Lys Leu Glu
260 265 270
Asp Ala Glu Lys Ala Ala Ser Thr Lys Glu Asp Tyr Arg Arg Glu
275 280 285
Arg Trp Arg Lys Pro Thr Tyr Ser Asp Lys Ala Gln Asn Cys Gln
290 295 300
Glu Ser Arg Glu Ser Asp Leu Val Lys Tyr Gly Asn Ser Ser Arg
305 310 315
Asp Arg Tyr Ala Thr Thr Asp Thr Ala Lys Asn Ser Asn Asn Glu
320 325 330
Lys Phe Ile Gly Asp Glu Lys Asp Lys Arg Pro Gly Ser Leu Glu
335 340 345
Thr Cys Arg Arg Glu Ser Asn Pro Arg Gln Asn Gln Glu Phe Ser
350 355 360
Phe Gly Asn Leu Arg Ala Lys Phe Leu Arg Pro Ser Asp Asp Glu

```
                    365                     370                     375
Glu Leu Ser Phe His Ser Lys Gly Arg Lys Phe Glu Pro Leu Ser
                380                 385                 390
Ser Ser Ser Ala Leu Val Ala Gln Gly Ser Leu Cys Ser Gly Phe
                395                 400                 405
Arg Lys Pro Thr Lys Asn Ser Glu Glu Arg Leu Thr Ser Trp Ser
                410                 415                 420
Arg Ser Asp Gly Arg Gly Asp Lys Lys His Ser Asn Gln Lys Pro
                425                 430                 435
Ser Glu Thr Ser Thr Asp Glu Tyr Gln His Val Pro Glu Asp Pro
                440                 445                 450
Arg Glu Lys Ser Gln Asp Glu Val Leu Arg Asp Asp Pro Pro Lys
                455                 460                 465
Lys Glu His Leu Arg Asp Thr Lys Ser Thr Phe Ala Gly Ser Pro
                470                 475                 480
Glu Arg Glu Ser Ile His Ile Leu Ser Val Asp Glu Lys Asn Lys
                485                 490                 495
Leu Gly Ala Lys Ile Ile Lys Ala Glu Met Met Gly Asn Met Glu
                500                 505                 510
Leu Ala Glu Gln Leu Lys Val Gln Leu Glu Lys Ala Asn Lys Phe
                515                 520                 525
Lys Glu Thr Ile Thr Gln Ile Pro Lys Lys Ser Gly Val Glu Asn
                530                 535                 540
Glu Asp Gln Gln Glu Val Ile Leu Val Arg Thr Asp Gln Ser Gly
                545                 550                 555
Arg Val Trp Pro Val Asn Thr Pro Gly Lys Ser Leu Glu Ser Gln
                560                 565                 570
Gly Gly Arg Arg Lys Arg Gln Met Val Ser Thr His Glu Glu Arg
                575                 580                 585
Glu Arg Val Arg Tyr Phe His Asp Asp Asp Asn Leu Ser Leu Asn
                590                 595                 600
Asp Leu Val Lys Asn Glu Lys Met Gly Thr Ala Glu Asn Gln Asn
                605                 610                 615
Lys Leu Phe Met Arg Met Ala Ser Lys Phe Met Gly Lys Thr Asp
                620                 625                 630
Gly Asp Tyr Tyr Thr Leu Asp Asp Met Phe Val Ser Lys Ala Ala
                635                 640                 645
Glu Arg Glu Arg Leu Gly Glu Glu Glu Glu Asn Gln Arg Lys Lys
                650                 655                 660
Ala Ile Ala Glu His Arg Ser Leu Ala Ala Gln Met Glu Lys Cys
                665                 670                 675
Leu Tyr Cys Phe Asp Ser Ser Gln Phe Pro Lys His Leu Ile Val
                680                 685                 690
Ala Ile Gly Val Lys Val Tyr Leu Cys Leu Pro Asn Val Arg Ser
                695                 700                 705
Leu Thr Glu Gly His Cys Leu Ile Val Pro Leu Gln His His Arg
                710                 715                 720
Ala Ala Thr Leu Leu Asp Glu Asp Ile Trp Glu Glu Ile Gln Met
                725                 730                 735
Phe Arg Lys Ser Leu Val Lys Met Phe Glu Asp Lys Gly Leu Asp
                740                 745                 750
Cys Ile Phe Leu Glu Thr Asn Met Ser Met Lys Lys Gln Tyr His
                755                 760                 765
Met Val Tyr Glu Cys Ile Pro Leu Pro Lys Glu Val Gly Asp Met
                770                 775                 780
Ala Pro Ile Tyr Phe Lys Lys Ala Ile Met Glu Ser Asp Glu Glu
                785                 790                 795
Trp Ser Met Asn Lys Lys Leu Met Asp Leu Ser Ser Lys Asp Ile
                800                 805                 810
Arg Lys Ser Val Pro Arg Gly Leu Pro Tyr Phe Ser Val Asp Phe
                815                 820                 825
Gly Leu His Gly Gly Phe Ala His Val Ile Glu Asp Gln His Lys
                830                 835                 840
```

```
Phe Pro His Tyr Phe Gly Lys Glu Ile Ile Gly Gly Met Leu Asp
                845                 850                 855
Ile Glu Pro Arg Leu Trp Arg Lys Gly Ile Arg Glu Ser Phe Glu
                860                 865                 870
Asp Gln Arg Lys Lys Ala Leu Gln Phe Ala Gln Trp Trp Lys Pro
                875                 880                 885
Tyr Asp Phe Thr Lys Ser Lys Asn Tyr
                890

<210> 34
<211> 653
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 368660CD1

<400> 34
Met Asp Arg Asp Leu Leu Arg Gln Ser Leu Asn Cys His Gly Ser
  1             5                   10                  15
Ser Leu Leu Ser Leu Leu Arg Ser Glu Gln Gln Asp Asn Pro His
                20                  25                  30
Phe Arg Ser Leu Leu Gly Ser Ala Ala Glu Pro Ala Arg Gly Pro
                35                  40                  45
Pro Pro Gln His Pro Leu Gln Gly Arg Lys Glu Lys Arg Val Asp
                50                  55                  60
Asn Ile Glu Ile Gln Lys Phe Ile Ser Lys Lys Ala Asp Leu Leu
                65                  70                  75
Phe Ala Leu Ser Trp Lys Ser Asp Ala Pro Ala Thr Ser Glu Ile
                80                  85                  90
Asn Glu Asp Ser Glu Asp His Tyr Ala Ile Met Pro Pro Leu Glu
                95                  100                 105
Gln Phe Met Glu Ile Pro Ser Met Asp Arg Arg Glu Leu Phe Phe
                110                 115                 120
Arg Asp Ile Glu Arg Gly Asp Ile Val Ile Gly Arg Ile Ser Ser
                125                 130                 135
Ile Arg Glu Phe Gly Phe Phe Met Val Leu Ile Cys Leu Gly Ser
                140                 145                 150
Gly Ile Met Arg Asp Ile Ala His Leu Glu Ile Thr Ala Leu Cys
                155                 160                 165
Pro Leu Arg Asp Val Pro Ser His Ser Asn His Gly Asp Pro Leu
                170                 175                 180
Ser Tyr Tyr Gln Thr Gly Asp Ile Ile Arg Ala Gly Ile Lys Asp
                185                 190                 195
Ile Asp Arg Tyr His Glu Lys Leu Ala Val Ser Leu Tyr Ser Ser
                200                 205                 210
Ser Leu Pro Pro His Leu Ser Gly Ile Lys Leu Gly Val Ile Ser
                215                 220                 225
Ser Glu Glu Leu Pro Leu Tyr Tyr Arg Arg Ser Val Glu Leu Asn
                230                 235                 240
Ser Asn Ser Leu Glu Ser Tyr Glu Asn Val Met Gln Ser Ser Leu
                245                 250                 255
Gly Phe Val Asn Pro Gly Val Val Glu Phe Leu Leu Glu Lys Leu
                260                 265                 270
Gly Ile Asp Glu Ser Asn Pro Pro Ser Leu Met Arg Gly Leu Gln
                275                 280                 285
Ser Lys Asn Phe Ser Glu Asp Asp Phe Ala Ser Ala Leu Arg Lys
                290                 295                 300
Lys Gln Ser Ala Ser Trp Ala Leu Lys Cys Val Lys Ile Gly Val
                305                 310                 315
Asp Tyr Phe Lys Val Gly Arg His Val Asp Ala Met Asn Glu Tyr
                320                 325                 330
```

```
Asn Lys Ala Leu Glu Ile Asp Lys Gln Asn Val Glu Ala Leu Val
                335                 340                 345
Ala Arg Gly Ala Leu Tyr Ala Thr Lys Gly Ser Leu Asn Lys Ala
                350                 355                 360
Ile Glu Asp Phe Glu Leu Ala Leu Glu Asn Cys Pro Thr His Arg
                365                 370                 375
Asn Ala Arg Lys Tyr Leu Cys Gln Thr Leu Val Glu Arg Gly Gly
                380                 385                 390
Gln Leu Glu Glu Glu Glu Lys Phe Leu Asn Ala Glu Ser Tyr Tyr
                395                 400                 405
Lys Lys Ala Leu Ala Leu Asp Glu Thr Phe Lys Asp Ala Glu Asp
                410                 415                 420
Ala Leu Gln Lys Leu His Lys Tyr Met Gln Lys Ser Leu Glu Leu
                425                 430                 435
Arg Glu Lys Gln Ala Glu Lys Glu Glu Lys Gln Lys Thr Lys Lys
                440                 445                 450
Ile Glu Thr Ser Ala Glu Lys Leu Arg Asn Val Leu Lys Glu Glu
                455                 460                 465
Lys Arg Leu Lys Lys Lys Arg Arg Lys Ser Thr Ser Ser Ser Ser
                470                 475                 480
Val Ser Ser Ala Asp Glu Ser Val Ser Ser Ser Ser Ser Ser Ser
                485                 490                 495
Ser Ser Gly His Lys Arg His Lys Lys His Lys Arg Asn Arg Ser
                500                 505                 510
Glu Ser Ser Arg Ser Ser Arg Arg His Ser Ser Arg Ala Ser Ser
                515                 520                 525
Asn Gln Ile Asp Gln Asn Arg Lys Asp Glu Cys Tyr Pro Val Pro
                530                 535                 540
Ala Asn Thr Ser Ala Ser Phe Leu Asn His Lys Gln Glu Val Glu
                545                 550                 555
Lys Leu Leu Gly Lys Gln Asp Arg Leu Gln Tyr Glu Lys Thr Gln
                560                 565                 570
Ile Lys Glu Lys Asp Arg Cys Pro Leu Ser Ser Ser Ser Leu Glu
                575                 580                 585
Ile Pro Asp Asp Phe Gly Val Tyr Ser Tyr Leu Phe Lys Lys Leu
                590                 595                 600
Thr Ile Lys Gln Pro Gln Ala Gly Pro Ser Gly Asp Ile Pro Glu
                605                 610                 615
Glu Gly Ile Val Ile Ile Asp Asp Ser Ser Ile His Val Thr Asp
                620                 625                 630
Pro Glu Asp Leu Gln Val Gly Gln Asp Met Glu Val Glu Asp Ser
                635                 640                 645
Gly Ile Asp Asp Pro Asp His Gly
                650

<210> 35
<211> 144
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2804990CD1

<400> 35
Met Leu Asn Arg Ile Ile Trp Leu Gln Ala Val Leu Glu Ile Ile
  1             5                   10                  15
Thr Asn Lys Thr Thr Gln Ala Leu Thr Val Leu Ala Trp Gln Glu
                20                  25                  30
Thr Leu Met Arg Asn Ala Ile Tyr Gln Asn Arg Leu Ala Leu Asp
                35                  40                  45
Tyr Leu Leu Ala Ala Glu Gly Gly Val Cys Glu Lys Phe Asp Leu
                50                  55                  60
```

Thr Asn Tyr Cys Leu His Ile Asp Asp Gln Gly Gln Val Val Glu
65 70 75
Asp Ile Val Lys Asp Ile Thr Lys Leu Ala His Ala Pro Val Gln
80 85 90
Val Trp His Gly Leu Asn Leu Gly Ala Met Phe Gly Asn Trp Phe
95 100 105
Pro Ala Ile Gly Gly Phe Lys Thr Leu Ile Ile Arg Val Ile Ile
110 115 120
Val Ile Gly Thr Cys Leu Leu Leu Pro Cys Leu Ile Pro Val Phe
125 130 135
Leu Gln Met Ile Lys Asn Phe Val Ala
140

<210> 36
<211> 424
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 168571CD1

<400> 36
Met Ser Pro Leu Cys Ser Leu Cys Gln Glu Gly Ser Trp Thr Gly
1 5 10 15
Pro Ala Ala Trp Met Pro Gly Pro Leu Gly Pro Glu His Gln Gly
20 25 30
Val Gln Pro Arg Thr Pro Gln Ala Trp Ala Pro Leu Pro Ala Glu
35 40 45
Gly Leu Trp Gly Ala Arg Gly Glu Ala Ser Arg His Gly Gly Cys
50 55 60
Pro Ser Pro Ser His Gly Leu Gly Pro His Ala Ala Leu Cys Leu
65 70 75
Pro Gln Glu Asn Pro Arg Leu Thr Glu Asp Phe Val Ser His Leu
80 85 90
Glu Thr Glu Leu Glu Gln Ser Arg Leu Arg Glu Thr Glu Thr Leu
95 100 105
Gly Ala Leu Arg Glu Met Gln Asp Lys Val Leu Asp Met Glu Lys
110 115 120
Arg Asn Ser Ser Leu Pro Asp Glu Asn Asn Val Ala Gln Leu Gln
125 130 135
Glu Glu Leu Lys Ala Leu Lys Val Arg Glu Gly Gln Ala Val Ala
140 145 150
Ser Thr Arg Glu Leu Lys Leu Gln Leu Gln Glu Leu Ser Asp Thr
155 160 165
Trp Gln Ala His Leu Ala Arg Gly Gly Arg Trp Lys Glu Ser Pro
170 175 180
Arg Lys Leu Val Val Gly Glu Leu Gln Asp Glu Leu Met Ser Val
185 190 195
Arg Leu Arg Glu Ala Gln Ala Leu Ala Glu Gly Arg Glu Leu Arg
200 205 210
Gln Arg Val Val Glu Leu Glu Thr Gln Asp His Ile His Arg Asn
215 220 225
Leu Leu Asn Arg Val Glu Ala Glu Arg Ala Ala Leu Gln Glu Lys
230 235 240
Leu Gln Tyr Leu Ala Ala Gln Asn Lys Gly Leu Gln Thr Gln Leu
245 250 255
Ser Glu Ser Arg Arg Lys Gln Ala Glu Ala Glu Cys Lys Ser Lys
260 265 270
Glu Glu Val Met Ala Val Arg Leu Arg Glu Ala Asp Ser Met Ala
275 280 285
Ala Val Ala Glu Met Arg Gln Arg Ile Ala Glu Leu Glu Ile Gln
290 295 300

```
Arg Glu Glu Gly Arg Ile Gln Gly Gln Leu Asn His Ser Asp Ser
                305                 310                 315
Ser Gln Tyr Ile Arg Glu Leu Lys Asp Gln Ile Glu Glu Leu Lys
                320                 325                 330
Ala Glu Val Arg Leu Leu Lys Gly Pro Pro Pro Phe Glu Asp Pro
                335                 340                 345
Leu Ala Phe Asp Gly Leu Ser Leu Ala Arg His Leu Asp Glu Asp
                350                 355                 360
Ser Leu Pro Ser Ser Asp Glu Glu Leu Leu Gly Val Gly Val Gly
                365                 370                 375
Ala Ala Leu Gln Asp Ala Leu Tyr Pro Leu Ser Pro Arg Asp Ala
                380                 385                 390
Arg Phe Phe Arg Arg Leu Glu Arg Pro Ala Lys Asp Ser Glu Gly
                395                 400                 405
Ser Ser Asp Ser Asp Ala Asp Glu Leu Ala Ala Pro Tyr Ser Gln
                410                 415                 420
Gly Leu Asp Asn


&lt;210&gt; 37
&lt;211&gt; 1351
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;220&gt;
&lt;221&gt; misc_feature
&lt;223&gt; Incyte ID No: 1286391CD1

&lt;400&gt; 37
Met Ala Ala Ala Val Pro Lys Ala Glu Asn Pro Ser Arg Thr Gln
 1               5                  10                  15
Val Pro Ser Ala Ala Pro Lys Leu Pro Thr Ser Arg Met Met Leu
                20                  25                  30
Ala Val His Thr Glu Pro Ala Ala Pro Glu Val Pro Leu Ala Pro
                35                  40                  45
Thr Lys Pro Thr Ala Gln Leu Met Ala Thr Ala Gln Lys Thr Val
                50                  55                  60
Val Asn Gln Pro Val Leu Val Ala Gln Val Glu Pro Thr Thr Pro
                65                  70                  75
Lys Thr Pro Gln Ala Gln Lys Met Pro Val Ala Lys Thr Ser Pro
                80                  85                  90
Ala Gly Pro Lys Thr Pro Lys Ala Gln Ala Gly Pro Ala Ala Thr
                95                 100                 105
Val Ser Lys Ala Pro Ala Ala Ser Lys Ala Pro Ala Ala Pro Lys
                110                 115                 120
Val Pro Val Thr Pro Arg Val Ser Arg Ala Pro Lys Thr Pro Ala
                125                 130                 135
Ala Gln Lys Val Pro Thr Asp Ala Gly Pro Thr Leu Asp Val Ala
                140                 145                 150
Arg Leu Leu Ser Glu Val Gln Pro Thr Ser Arg Ala Ser Val Ser
                155                 160                 165
Leu Leu Lys Gly Gln Gly Gln Ala Gly Arg Gln Gly Pro Gln Ser
                170                 175                 180
Ser Gly Thr Leu Ala Leu Ser Ser Lys His Gln Phe Gln Met Glu
                185                 190                 195
Gly Leu Leu Gly Ala Trp Glu Gly Ala Pro Arg Gln Pro Pro Arg
                200                 205                 210
His Leu Gln Ala Asn Ser Thr Val Thr Ser Phe Gln Arg Tyr His
                215                 220                 225
Glu Ala Leu Asn Thr Pro Phe Glu Leu Asn Leu Ser Gly Glu Pro
                230                 235                 240
Gly Asn Gln Gly Leu Arg Arg Val Val Ile Asp Gly Ser Ser Val
                245                 250                 255
```

Ala Met Val His Gly Leu Gln His Phe Phe Ser Cys Arg Gly Ile
260 265 270
Ala Met Ala Val Gln Phe Phe Trp Asn Arg Gly His Arg Glu Val
275 280 285
Thr Val Phe Val Pro Thr Trp Gln Leu Lys Lys Asn Arg Arg Val
290 295 300
Arg Glu Ser His Phe Leu Thr Lys Leu His Ser Leu Lys Met Leu
305 310 315
Ser Ile Thr Pro Ser Gln Leu Glu Asn Gly Lys Lys Ile Thr Thr
320 325 330
Tyr Asp Tyr Arg Phe Met Val Lys Leu Ala Glu Glu Thr Asp Gly
335 340 345
Ile Ile Val Thr Asn Glu Gln Ile His Ile Leu Met Asn Ser Ser
350 355 360
Lys Lys Leu Met Val Lys Asp Arg Leu Leu Pro Phe Thr Phe Ala
365 370 375
Gly Asn Leu Phe Met Val Pro Asp Asp Pro Leu Gly Arg Asp Gly
380 385 390
Pro Thr Leu Asp Glu Phe Leu Lys Lys Pro Asn Arg Leu Asp Thr
395 400 405
Asp Ile Gly Asn Phe Leu Lys Val Trp Lys Thr Leu Pro Pro Ser
410 415 420
Ser Ala Ser Val Thr Glu Leu Ser Asp Asp Ala Asp Ser Gly Pro
425 430 435
Leu Glu Ser Leu Pro Asn Met Glu Glu Val Arg Glu Glu Lys Glu
440 445 450
Glu Arg Gln Asp Glu Glu Gln Arg Gln Gly Gln Gly Thr Gln Lys
455 460 465
Ala Ala Glu Glu Asp Asp Leu Asp Ser Ser Leu Ala Ser Val Phe
470 475 480
Arg Val Glu Cys Pro Ser Leu Ser Glu Glu Ile Leu Arg Cys Leu
485 490 495
Ser Leu His Asp Pro Pro Asp Gly Ala Leu Asp Ile Asp Leu Leu
500 505 510
Pro Gly Ala Ala Ser Pro Tyr Leu Gly Ile Pro Trp Asp Gly Lys
515 520 525
Ala Pro Cys Gln Gln Val Leu Ala His Leu Ala Gln Leu Thr Ile
530 535 540
Pro Ser Asn Phe Thr Ala Leu Ser Phe Phe Met Gly Phe Met Asp
545 550 555
Ser His Arg Asp Ala Ile Pro Asp Tyr Glu Ala Leu Val Gly Pro
560 565 570
Leu His Ser Leu Leu Lys Gln Lys Pro Asp Trp Gln Trp Asp Gln
575 580 585
Glu His Glu Glu Ala Phe Leu Ala Leu Lys Arg Ala Leu Val Ser
590 595 600
Ala Leu Cys Leu Met Ala Pro Asn Ser Gln Leu Pro Phe Arg Leu
605 610 615
Glu Val Thr Val Ser His Val Ala Leu Thr Ala Ile Leu His Gln
620 625 630
Glu His Ser Gly Arg Lys His Pro Ile Ala Tyr Thr Ser Lys Pro
635 640 645
Leu Leu Pro Asp Glu Glu Ser Gln Gly Pro Gln Ser Gly Gly Asp
650 655 660
Ser Pro Tyr Ala Val Ala Trp Ala Leu Lys His Phe Ser Arg Cys
665 670 675
Ile Gly Asp Thr Pro Val Val Leu Asp Leu Ser Tyr Ala Ser Arg
680 685 690
Thr Thr Ala Asp Pro Glu Val Arg Glu Gly Arg Arg Val Ser Lys
695 700 705
Ala Trp Leu Ile Arg Trp Ser Leu Leu Val Gln Asp Lys Gly Lys
710 715 720
Arg Ala Leu Glu Leu Ala Leu Leu Gln Gly Leu Leu Gly Glu Asn

725 730 735
Arg Leu Leu Thr Pro Ala Ala Ser Met Pro Arg Phe Phe Gln Val
740 745 750
Leu Pro Pro Phe Ser Asp Leu Ser Thr Phe Val Cys Ile His Met
755 760 765
Ser Gly Tyr Cys Phe Tyr Arg Glu Asp Glu Trp Cys Ala Gly Phe
770 775 780
Gly Leu Tyr Val Leu Ser Pro Thr Ser Pro Pro Val Ser Leu Ser
785 790 795
Phe Ser Cys Ser Pro Tyr Thr Pro Thr Tyr Ala His Leu Ala Ala
800 805 810
Val Ala Cys Gly Leu Glu Arg Phe Gly Gln Ser Pro Leu Pro Val
815 820 825
Val Phe Leu Thr His Cys Asn Trp Ile Phe Ser Leu Leu Trp Glu
830 835 840
Leu Leu Pro Leu Trp Arg Ala Arg Gly Phe Leu Ser Ser Asp Gly
845 850 855
Ala Pro Leu Pro His Pro Ser Leu Leu Ser Tyr Ile Ile Ser Leu
860 865 870
Thr Ser Gly Leu Ser Ser Leu Pro Phe Ile Tyr Arg Thr Ser Tyr
875 880 885
Arg Gly Ser Leu Phe Ala Val Thr Val Asp Thr Leu Ala Lys Gln
890 895 900
Gly Ala Gln Gly Gly Gly Gln Trp Trp Ser Leu Pro Lys Asp Val
905 910 915
Pro Ala Pro Thr Val Ser Pro His Ala Met Gly Lys Arg Pro Asn
920 925 930
Leu Leu Ala Leu Gln Leu Ser Asp Ser Thr Leu Ala Asp Ile Ile
935 940 945
Ala Arg Leu Gln Ala Gly Gln Lys Leu Ser Gly Ser Ser Pro Phe
950 955 960
Ser Ser Ala Phe Asn Ser Leu Ser Leu Asp Lys Glu Ser Gly Leu
965 970 975
Leu Met Phe Lys Gly Asp Lys Lys Pro Arg Val Trp Val Val Pro
980 985 990
Thr Gln Leu Arg Arg Asp Leu Ile Phe Ser Val His Asp Ile Pro
995 1000 1005
Leu Gly Ala His Gln Arg Pro Glu Glu Thr Tyr Lys Lys Leu Arg
1010 1015 1020
Leu Leu Gly Trp Trp Pro Gly Met Gln Glu His Val Lys Asp Tyr
1025 1030 1035
Cys Arg Ser Cys Leu Phe Cys Ile Pro Arg Asn Leu Ile Gly Ser
1040 1045 1050
Glu Leu Lys Val Ile Glu Ser Pro Trp Pro Leu Arg Ser Thr Ala
1055 1060 1065
Pro Trp Ser Asn Leu Gln Ile Glu Val Val Gly Pro Val Thr Ile
1070 1075 1080
Ser Glu Glu Gly His Lys His Val Leu Ile Val Ala Asp Pro Asn
1085 1090 1095
Thr Arg Trp Val Glu Ala Phe Pro Leu Lys Pro Tyr Thr His Thr
1100 1105 1110
Ala Val Ala Gln Val Leu Leu Gln His Val Phe Ala Arg Trp Gly
1115 1120 1125
Val Pro Val Arg Leu Glu Ala Ala Gln Gly Pro Gln Phe Ala Arg
1130 1135 1140
His Val Leu Val Ser Cys Gly Leu Ala Leu Gly Ala Gln Val Ala
1145 1150 1155
Ser Leu Ser Arg Asp Leu Gln Phe Pro Cys Leu Thr Ser Ser Gly
1160 1165 1170
Ala Tyr Trp Glu Phe Lys Arg Ala Leu Lys Glu Phe Ile Phe Leu
1175 1180 1185
His Gly Lys Lys Trp Ala Ala Ser Leu Pro Leu Leu His Leu Ala
1190 1195 1200

Phe Arg Ala Ser Ser Thr Asp Ala Thr Pro Phe Lys Val Leu Thr
1205 1210 1215
Gly Gly Glu Ser Arg Leu Thr Glu Pro Leu Trp Trp Glu Met Ser
1220 1225 1230
Ser Ala Asn Ile Glu Gly Leu Lys Met Asp Val Phe Leu Leu Gln
1235 1240 1245
Leu Val Gly Glu Leu Leu Glu Leu His Trp Arg Val Ala Asp Lys
1250 1255 1260
Ala Ser Glu Lys Ala Glu Asn Arg Arg Phe Lys Arg Glu Ser Gln
1265 1270 1275
Glu Lys Glu Trp Asn Val Gly Asp Gln Val Leu Leu Leu Ser Leu
1280 1285 1290
Pro Arg Asn Gly Ser Ser Ala Lys Trp Val Gly Pro Phe Tyr Ile
1295 1300 1305
Gly Asp Arg Leu Ser Leu Ser Leu Tyr Arg Ile Trp Gly Phe Pro
1310 1315 1320
Thr Pro Glu Lys Leu Gly Cys Ile Tyr Pro Ser Ser Leu Met Lys
1325 1330 1335
Ala Phe Ala Lys Ser Gly Thr Pro Leu Ser Phe Lys Val Leu Glu
1340 1345 1350
Gln

<210> 38
<211> 78
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2007684CD1

<400> 38
Met Gln Glu Met Val Arg Glu Leu Trp Met Trp Asn Val Glu Glu
1 5 10 15
Glu Glu His Glu Val Gly Ile Cys Thr Trp Gly Gly Gln His Cys
20 25 30
Gly Cys Pro Ala Lys Ser Leu Pro Gly Pro His Pro Gly Gly Val
35 40 45
Ser Ala Pro Gln Ser Ala Ser Gln Leu Met Val Lys Leu Leu Val
50 55 60
Trp Gln Lys Ser Val His Lys Leu Arg Lys Leu Leu Glu Lys Thr
65 70 75
Glu Asn Tyr

<210> 39
<211> 411
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2227040CD1

<400> 39
Met Thr Glu Met Ser Glu Lys Glu Asn Glu Pro Asp Asp Ala Ala
1 5 10 15
Thr His Ser Pro Pro Gly Thr Val Ser Ala Leu Gln Glu Thr Lys
20 25 30
Leu Gln Arg Phe Lys Arg Ser Leu Ser Leu Lys Thr Ile Leu Arg
35 40 45
Ser Lys Ser Leu Glu Asn Phe Phe Leu Arg Ser Gly Ser Glu Leu

```
                    50                          55                          60
Lys Cys Pro Thr Glu Val Leu Leu Thr Pro Pro Thr Pro Leu Pro
                    65                          70                          75
Pro Pro Ser Pro Pro Pro Thr Ala Ser Asp Arg Gly Leu Ala Thr
                    80                          85                          90
Pro Ser Pro Ser Pro Cys Pro Val Pro Arg Pro Leu Ala Ala Leu
                    95                         100                         105
Lys Pro Val Thr Leu His Ser Phe Gln Glu His Val Phe Lys Arg
                   110                         115                         120
Ala Ser Pro Cys Glu Leu Cys His Gln Leu Ile Val Gly Asn Ser
                   125                         130                         135
Lys Gln Gly Leu Arg Cys Lys Met Cys Lys Val Ser Val His Leu
                   140                         145                         150
Trp Cys Ser Glu Glu Ile Ser His Gln Gln Cys Pro Gly Lys Thr
                   155                         160                         165
Ser Thr Ser Phe Arg Arg Asn Phe Ser Ser Pro Leu Leu Val His
                   170                         175                         180
Glu Pro Pro Pro Val Cys Ala Thr Ser Lys Glu Ser Pro Pro Thr
                   185                         190                         195
Gly Asp Ser Gly Lys Val Asp Pro Val Tyr Glu Thr Leu Arg Tyr
                   200                         205                         210
Gly Thr Ser Leu Ala Leu Met Asn Arg Ser Ser Phe Ser Ser Thr
                   215                         220                         225
Ser Glu Ser Pro Thr Arg Ser Leu Ser Glu Arg Asp Glu Leu Thr
                   230                         235                         240
Glu Asp Gly Glu Gly Ser Ile Arg Ser Ser Glu Glu Gly Pro Gly
                   245                         250                         255
Asp Ser Ala Ser Pro Val Phe Thr Ala Pro Ala Glu Ser Glu Gly
                   260                         265                         270
Pro Gly Pro Glu Glu Lys Ser Pro Gly Gln Gln Leu Pro Lys Ala
                   275                         280                         285
Thr Leu Arg Lys Asp Val Gly Pro Met Tyr Ser Tyr Val Ala Leu
                   290                         295                         300
Tyr Lys Phe Leu Pro Gln Glu Asn Asn Asp Leu Ala Leu Gln Pro
                   305                         310                         315
Gly Asp Arg Ile Met Leu Val Asp Asp Ser Asn Glu Asp Trp Trp
                   320                         325                         330
Lys Gly Lys Ile Gly Asp Arg Val Gly Phe Phe Pro Ala Asn Phe
                   335                         340                         345
Val Gln Arg Val Arg Pro Gly Glu Asn Val Trp Arg Cys Cys Gln
                   350                         355                         360
Pro Phe Ser Gly Asn Lys Glu Gln Gly Tyr Met Ser Leu Lys Glu
                   365                         370                         375
Asn Gln Ile Cys Val Gly Val Gly Arg Ser Lys Asp Ala Asp Gly
                   380                         385                         390
Phe Ile Arg Val Ser Ser Gly Lys Lys Arg Gly Leu Val Pro Val
                   395                         400                         405
Asp Ala Leu Thr Glu Ile
                   410

<210> 40
<211> 1704
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4346130CD1

<400> 40
Met Ser Ser Val Ser Glu Val Asn Val Asp Ile Lys Asp Phe Leu
  1               5                  10                  15
Met Ser Ile Asn Leu Glu Gln Tyr Leu Leu His Phe His Glu Ser
```

```
                20                  25                  30
Gly Phe Thr Thr Val Lys Asp Cys Ala Ala Ile Asn Asp Ser Leu
                35                  40                  45
Leu Gln Lys Ile Gly Ile Ser Pro Thr Gly His Arg Arg Arg Ile
                50                  55                  60
Leu Lys Gln Leu Gln Ile Ile Leu Ser Lys Met Gln Asp Ile Pro
                65                  70                  75
Ile Tyr Ala Asn Val His Lys Thr Lys Lys Asn Asp Asp Pro Ser
                80                  85                  90
Lys Asp Tyr His Val Pro Ser Ser Asp Gln Asn Ile Cys Ile Glu
                95                  100                 105
Leu Ser Asn Ser Gly Ser Val Gln Thr Ser Ser Pro Pro Gln Leu
                110                 115                 120
Glu Thr Val Arg Lys Asn Leu Glu Asp Ser Asp Ala Ser Val Glu
                125                 130                 135
Arg Ser Gln Tyr Pro Gln Ser Asp Asp Lys Leu Ser Pro Pro Lys
                140                 145                 150
Arg Asp Phe Pro Thr Ala Glu Glu Pro His Leu Asn Leu Gly Ser
                155                 160                 165
Leu Asn Asp Ser Leu Phe Gly Ser Asp Asn Ile Lys Ile Glu Ser
                170                 175                 180
Leu Ile Thr Lys Lys Thr Val Asp His Thr Val Glu Glu Gln Gln
                185                 190                 195
Thr Glu Lys Val Lys Leu Ile Thr Glu Asn Leu Ser Lys Leu Pro
                200                 205                 210
Asn Ala Asp Ser Glu Cys Leu Ser Phe Val Gly Cys Ser Thr Ser
                215                 220                 225
Gly Thr Asn Ser Gly Asn Gly Thr Asn Gly Leu Leu Glu Gly Ser
                230                 235                 240
Pro Pro Ser Pro Phe Phe Lys Phe Gln Gly Glu Met Ile Val Asn
                245                 250                 255
Asp Leu Tyr Val Pro Ser Ser Pro Ile Leu Ala Pro Val Arg Ser
                260                 265                 270
Arg Ser Lys Leu Val Ser Arg Pro Ser Arg Ser Phe Leu Leu Arg
                275                 280                 285
His Arg Pro Val Pro Glu Ile Pro Gly Ser Thr Lys Gly Val Ser
                290                 295                 300
Gly Ser Tyr Phe Arg Glu Arg Arg Asn Val Ala Thr Ser Thr Glu
                305                 310                 315
Lys Ser Val Ala Trp Gln Asn Ser Asn Glu Glu Asn Ser Ser Ser
                320                 325                 330
Ile Phe Pro Tyr Gly Glu Thr Phe Leu Phe Gln Arg Leu Glu Asn
                335                 340                 345
Ser Lys Lys Arg Ser Ile Lys Asn Glu Phe Leu Thr Gln Gly Glu
                350                 355                 360
Ala Leu Lys Gly Glu Ala Ala Thr Ala Thr Asn Ser Phe Ile Ile
                365                 370                 375
Lys Ser Ser Ile Tyr Asp Asn Arg Lys Glu Lys Ile Ser Glu Asp
                380                 385                 390
Lys Val Glu Asp Ile Trp Ile Pro Arg Glu Asp Lys Asn Asn Phe
                395                 400                 405
Leu Ile Asp Thr Ala Ser Glu Ser Glu Tyr Ser Thr Val Glu Glu
                410                 415                 420
Cys Phe Gln Ser Leu Arg Arg Lys Asn Ser Lys Ala Ser Lys Ser
                425                 430                 435
Arg Thr Gln Lys Ala Leu Ile Leu Asp Ser Val Asn Arg His Ser
                440                 445                 450
Tyr Pro Leu Ser Ser Thr Ser Gly Asn Ala Asp Ser Ser Ala Val
                455                 460                 465
Ser Ser Gln Ala Ile Ser Pro Tyr Ala Cys Phe Tyr Gly Ala Ser
                470                 475                 480
Ala Lys Lys Val Lys Ser Gly Trp Leu Asp Lys Leu Ser Pro Gln
                485                 490                 495
```

Gly Lys Arg Met Phe Gln Lys Arg Trp Val Lys Phe Asp Gly Leu
500 505 510
Ser Ile Ser Tyr Tyr Asn Asn Glu Lys Glu Met Tyr Ser Lys Gly
515 520 525
Ile Ile Pro Leu Ser Ala Ile Ser Thr Val Arg Val Gln Gly Asp
530 535 540
Asn Lys Phe Glu Val Val Thr Thr Gln Arg Thr Phe Val Phe Arg
545 550 555
Val Glu Lys Glu Glu Glu Arg Asn Asp Trp Ile Ser Ile Leu Leu
560 565 570
Asn Ala Leu Lys Ser Gln Ser Leu Thr Ser Gln Ser Gln Ala Val
575 580 585
Val Thr Pro Glu Lys Cys Gly Tyr Leu Glu Leu Arg Gly Tyr Lys
590 595 600
Ala Lys Ile Phe Thr Val Leu Ser Gly Asn Ser Val Trp Leu Cys
605 610 615
Lys Asn Glu Gln Asp Phe Lys Ser Gly Leu Gly Ile Thr Ile Ile
620 625 630
Pro Met Asn Val Ala Asn Val Lys Gln Val Asp Arg Thr Val Lys
635 640 645
Gln Ser Phe Glu Ile Ile Thr Pro Tyr Arg Ser Phe Ser Phe Thr
650 655 660
Ala Glu Thr Glu Lys Glu Lys Gln Asp Trp Ile Glu Ala Val Gln
665 670 675
Gln Ser Ile Ala Glu Thr Leu Ser Asp Tyr Glu Val Ala Glu Lys
680 685 690
Ile Trp Phe Asn Glu Ser Asn Arg Ser Cys Ala Asp Cys Lys Ala
695 700 705
Pro Asp Pro Asp Trp Ala Ser Ile Asn Leu Cys Val Val Ile Cys
710 715 720
Lys Lys Cys Ala Gly Gln His Arg Ser Leu Gly Pro Lys Asp Ser
725 730 735
Lys Val Arg Ser Leu Lys Met Asp Ala Ser Ile Trp Ser Asn Glu
740 745 750
Leu Ile Glu Leu Phe Ile Val Ile Gly Asn Lys Arg Ala Asn Asp
755 760 765
Phe Trp Ala Gly Asn Leu Gln Lys Asp Glu Glu Leu His Met Asp
770 775 780
Ser Pro Val Glu Lys Arg Lys Asn Phe Ile Thr Gln Lys Tyr Lys
785 790 795
Glu Gly Lys Phe Arg Lys Thr Leu Leu Ala Ser Leu Thr Lys Glu
800 805 810
Glu Leu Asn Lys Ala Leu Cys Ala Ala Val Val Lys Pro Asp Val
815 820 825
Leu Glu Thr Met Ala Leu Leu Phe Ser Gly Ala Asp Val Met Cys
830 835 840
Ala Thr Gly Asp Pro Val His Ser Thr Pro Tyr Leu Leu Ala Lys
845 850 855
Lys Ala Gly Gln Ser Leu Gln Met Glu Phe Leu Tyr His Asn Lys
860 865 870
Phe Ser Asp Phe Pro Gln His Asp Ile His Ser Glu Gly Val Leu
875 880 885
Ser Gln Glu Ser Ser Gln Ser Thr Phe Leu Cys Asp Phe Leu Tyr
890 895 900
Gln Ala Pro Ser Ala Ala Ser Lys Leu Ser Ser Glu Lys Lys Leu
905 910 915
Leu Glu Glu Thr Asn Lys Lys Trp Cys Val Leu Glu Gly Gly Phe
920 925 930
Leu Ser Tyr Tyr Glu Asn Asp Lys Ser Thr Thr Pro Asn Gly Thr
935 940 945
Ile Asn Ile Asn Glu Val Ile Cys Leu Ala Ile His Lys Glu Asp
950 955 960
Phe Tyr Leu Asn Thr Gly Pro Ile Phe Ile Phe Glu Ile Tyr Leu

965 970 975
Pro Ser Glu Arg Val Phe Leu Phe Gly Ala Glu Thr Ser Gln Ala
980 985 990
Gln Arg Lys Trp Thr Glu Ala Ile Ala Lys His Phe Val Pro Leu
995 1000 1005
Phe Ala Glu Asn Leu Thr Glu Ala Asp Tyr Asp Leu Ile Gly Gln
1010 1015 1020
Leu Phe Tyr Lys Asp Cys His Ala Leu Asp Gln Trp Arg Lys Gly
1025 1030 1035
Trp Phe Ala Met Asp Lys Ser Ser Leu His Phe Cys Leu Gln Met
1040 1045 1050
Gln Glu Val Gln Gly Asp Arg Met His Leu Arg Arg Leu Gln Glu
1055 1060 1065
Leu Thr Ile Ser Thr Met Val Gln Asn Gly Glu Lys Leu Asp Val
1070 1075 1080
Leu Leu Leu Val Glu Lys Gly Arg Thr Leu Tyr Ile His Gly His
1085 1090 1095
Thr Lys Leu Asp Phe Thr Val Trp His Thr Ala Ile Glu Lys Ala
1100 1105 1110
Ala Gly Thr Asp Gly Asn Ala Leu Gln Asp Gln Gln Leu Ser Lys
1115 1120 1125
Asn Asp Val Pro Ile Ile Val Asn Ser Cys Ile Ala Phe Val Thr
1130 1135 1140
Gln Tyr Gly Leu Gly Cys Lys Tyr Ile Tyr Gln Lys Asn Gly Asp
1145 1150 1155
Pro Leu His Ile Ser Glu Leu Leu Glu Ser Phe Lys Lys Asp Ala
1160 1165 1170
Arg Ser Phe Lys Leu Arg Ala Gly Lys His Gln Leu Glu Asp Val
1175 1180 1185
Thr Ala Val Leu Lys Ser Phe Leu Ser Asp Ile Asp Asp Ala Leu
1190 1195 1200
Leu Thr Lys Glu Leu Tyr Pro Tyr Trp Ile Ser Ala Leu Asp Thr
1205 1210 1215
Gln Asp Asp Lys Glu Arg Ile Lys Lys Tyr Gly Ala Phe Ile Arg
1220 1225 1230
Ser Leu Pro Gly Val Asn Arg Ala Thr Leu Ala Ala Ile Ile Glu
1235 1240 1245
His Leu Tyr Arg Val Gln Lys Cys Ser Glu Ile Asn His Met Asn
1250 1255 1260
Ala His Asn Leu Ala Leu Val Phe Ser Ser Cys Leu Phe Gln Thr
1265 1270 1275
Lys Gly Gln Thr Ser Glu Glu Val Asn Val Ile Glu Asp Leu Ile
1280 1285 1290
Asn Asn Tyr Val Glu Ile Phe Glu Val Lys Glu Asp Gln Val Lys
1295 1300 1305
Gln Met Asp Ile Glu Asn Ser Phe Ile Thr Lys Trp Lys Asp Thr
1310 1315 1320
Gln Val Ser Gln Ala Gly Asp Leu Leu Ile Glu Val Tyr Val Glu
1325 1330 1335
Arg Lys Glu Pro Asp Cys Ser Ile Ile Ile Arg Ile Ser Pro Val
1340 1345 1350
Met Glu Ala Glu Glu Leu Thr Asn Asp Ile Leu Ala Ile Lys Asn
1355 1360 1365
Ile Ile Pro Thr Lys Gly Asp Ile Trp Ala Thr Phe Glu Val Ile
1370 1375 1380
Glu Asn Glu Glu Leu Glu Arg Pro Leu His Tyr Lys Glu Asn Val
1385 1390 1395
Leu Glu Gln Val Leu Arg Trp Ser Ser Leu Ala Glu Pro Gly Ser
1400 1405 1410
Ala Tyr Leu Val Val Lys Arg Phe Leu Thr Ala Asp Thr Ile Lys
1415 1420 1425
His Cys Ser Asp Arg Ser Thr Leu Gly Ser Ile Lys Glu Gly Ile
1430 1435 1440

Leu Lys Ile Lys Glu Glu Pro Ser Lys Ile Leu Ser Gly Asn Lys
1445 1450 1455
Phe Gln Asp Arg Tyr Phe Val Leu Arg Asp Gly Phe Leu Phe Leu
1460 1465 1470
Tyr Lys Asp Val Lys Ser Ser Lys His Asp Lys Met Phe Ser Leu
1475 1480 1485
Ser Ser Met Lys Phe Tyr Arg Gly Val Lys Lys Lys Met Lys Pro
1490 1495 1500
Pro Thr Ser Trp Gly Leu Thr Ala Tyr Ser Glu Lys His His Trp
1505 1510 1515
His Leu Cys Cys Asp Ser Ser Gln Thr Gln Thr Glu Trp Met Thr
1520 1525 1530
Ser Ile Phe Ile Ala Gln His Glu Tyr Asp Ile Trp Pro Pro Ala
1535 1540 1545
Gly Lys Glu Arg Lys Arg Ser Ile Thr Lys Asn Pro Lys Ile Gly
1550 1555 1560
Gly Leu Pro Leu Ile Pro Ile Gln His Glu Gly Asn Ala Thr Leu
1565 1570 1575
Ala Arg Lys Asn Ile Glu Ser Ala Arg Ala Glu Leu Glu Arg Leu
1580 1585 1590
Arg Leu Ser Glu Lys Cys Asp Lys Glu Ser Val Asp Ser Ser Leu
1595 1600 1605
Lys Glu Arg Ala Ser Met Val Ala His Cys Leu Glu His Lys Asp
1610 1615 1620
Asp Lys Leu Arg Asn Arg Pro Arg Lys His Arg Ser Phe Asn Cys
1625 1630 1635
Leu Glu Asp Thr Glu Pro Glu Ala Pro Leu Gly Gln Pro Lys Gly
1640 1645 1650
His Lys Gly Leu Lys Thr Leu Arg Lys Thr Glu Asp Arg Asn Ser
1655 1660 1665
Lys Ala Thr Leu Asp Ser Asp His Lys Leu Pro Ser Arg Val Ile
1670 1675 1680
Glu Glu Leu Asn Val Val Leu Gln Arg Ser Arg Thr Leu Pro Lys
1685 1690 1695
Glu Leu Gln Asp Glu Gln Ile Leu Lys
1700

<210> 41
<211> 243
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 55117040CD1

<400> 41
Met Val Ala Glu Val Asp Ser Met Pro Ala Ala Ser Ser Val Lys
1 5 10 15
Lys Pro Phe Gly Leu Arg Ser Lys Met Gly Lys Trp Cys Cys His
20 25 30
Cys Phe Pro Cys Cys Arg Gly Ser Gly Lys Ser Asn Val Gly Thr
35 40 45
Ser Gly Asp His Asn Asp Ser Ser Val Lys Thr Leu Gly Ser Lys
50 55 60
Arg Cys Lys Trp Cys Cys His Cys Phe Pro Cys Cys Arg Gly Ser
65 70 75
Gly Lys Ser Asn Val Val Ala Trp Gly Asp Tyr Asp Asp Ser Ala
80 85 90
Phe Met Asp Pro Arg Tyr His Val His Gly Glu Asp Leu Asp Lys
95 100 105
Leu His Arg Ala Ala Trp Trp Gly Lys Val Pro Arg Lys Asp Leu
110 115 120

```
Ile Val Met Leu Arg Asp Thr Asp Val Asn Lys Arg Asp Lys Gln
                125                 130                 135
Lys Arg Thr Ala Leu His Leu Ala Ser Ala Asn Gly Asn Ser Glu
                140                 145                 150
Val Val Lys Leu Val Leu Asp Arg Arg Cys Gln Leu Asn Val Leu
                155                 160                 165
Asp Asn Lys Lys Arg Thr Ala Leu Thr Lys Ala Val Gln Cys Gln
                170                 175                 180
Glu Asp Glu Cys Ala Leu Met Leu Leu Glu His Gly Thr Asp Pro
                185                 190                 195
Asn Ile Pro Asp Glu Tyr Gly Asn Thr Thr Leu His Tyr Ala Val
                200                 205                 210
Tyr Asn Glu Asp Lys Leu Met Ala Lys His Cys Ser Tyr Thr Val
                215                 220                 225
Leu Ile Ser Asn Gln Lys Thr Ala Trp Pro His Thr Thr Ala Thr
                230                 235                 240
Trp Tyr Thr


&lt;210&gt; 42
&lt;211&gt; 248
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;220&gt;
&lt;221&gt; misc_feature
&lt;223&gt; Incyte ID No: 7472392CD1

&lt;400&gt; 42
Met Asp Val Leu His Ala Ser Val Arg Arg Ser Thr Ile Val Cys
 1               5                  10                  15
Met Glu Glu Thr Glu Phe Leu Val Val Asp Arg Glu Asp Phe Phe
                20                  25                  30
Ala Asn Lys Leu Asp Gln Glu Val Gln Lys Asp Ala Gln Tyr Arg
                35                  40                  45
Phe Glu Phe Phe Arg Lys Met Glu Leu Phe Ala Ser Trp Ser Asp
                50                  55                  60
Glu Lys Leu Trp Gln Leu Val Ala Met Ala Lys Ile Glu Arg Phe
                65                  70                  75
Ser Tyr Gly Gln Leu Ile Ser Lys Asp Phe Gly Glu Ser Pro Phe
                80                  85                  90
Ile Met Phe Ile Ser Lys Gly Ser Cys Glu Val Leu Arg Leu Leu
                95                  100                 105
Asp Leu Gly Ala Ser Pro Ser Tyr Arg Arg Trp Ile Trp Gln His
                110                 115                 120
Leu Glu Leu Ile Asp Gly Arg Pro Leu Lys Thr His Leu Ser Glu
                125                 130                 135
Tyr Ser Pro Met Glu Arg Phe Lys Glu Phe Gln Ile Lys Ser Tyr
                140                 145                 150
Pro Leu Gln Asp Phe Ser Ser Leu Lys Leu Pro His Leu Lys Lys
                155                 160                 165
Ala Trp Gly Leu Gln Gly Thr Ser Phe Ser Arg Lys Ile Arg Thr
                170                 175                 180
Ser Gly Asp Thr Leu Pro Lys Met Leu Gly Pro Lys Ile Gln Ser
                185                 190                 195
Arg Pro Ala Gln Ser Ile Lys Cys Ala Met Ile Asn Ile Lys Pro
                200                 205                 210
Gly Glu Leu Pro Lys Glu Ala Ala Val Gly Ala Tyr Val Lys Val
                215                 220                 225
His Thr Val Glu Gln Gly Glu Ile Leu Val Ser Val Pro Arg Ala
                230                 235                 240
Leu Phe Thr Met Glu Tyr Val Thr
                245
```

<210> 43
<211> 310
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4028960CD1

<400> 43
```
Met Gly Lys Arg Arg Cys Val Pro Pro Leu Glu Pro Lys Leu Ala
  1               5                  10                  15
Ala Gly Cys Cys Gly Val Lys Lys Pro Lys Leu Ser Gly Ser Gly
                 20                  25                  30
Thr His Ser His Gly Asn Gln Ser Thr Thr Val Pro Gly Ser Ser
                 35                  40                  45
Ser Gly Pro Leu Gln Asn His Gln His Val Asp Ser Ser Ser Gly
                 50                  55                  60
Arg Glu Asn Val Ser Asp Leu Thr Leu Gly Pro Gly Asn Ser Pro
                 65                  70                  75
Ile Thr Arg Met Asn Pro Ala Ser Gly Ala Leu Ser Pro Leu Pro
                 80                  85                  90
Arg Pro Asn Gly Thr Ala Asn Thr Thr Lys Asn Leu Val Val Thr
                 95                 100                 105
Ala Glu Met Cys Cys Tyr Cys Phe Asp Val Leu Tyr Cys His Leu
                110                 115                 120
Tyr Gly Phe Pro Gln Pro Arg Leu Pro Arg Phe Thr Asn Asp Pro
                125                 130                 135
Tyr Pro Leu Phe Val Thr Trp Lys Thr Gly Arg Asp Lys Arg Leu
                140                 145                 150
Arg Gly Cys Ile Gly Thr Phe Ser Ala Met Asn Leu His Ser Gly
                155                 160                 165
Leu Arg Glu Tyr Thr Leu Thr Ser Ala Leu Lys Asp Ser Arg Phe
                170                 175                 180
Pro Pro Leu Thr Arg Glu Glu Leu Pro Lys Leu Phe Cys Ser Val
                185                 190                 195
Ser Leu Leu Thr Asn Phe Glu Asp Ala Ser Asp Tyr Leu Asp Trp
                200                 205                 210
Glu Val Gly Val His Gly Ile Arg Ile Glu Phe Ile Asn Glu Lys
                215                 220                 225
Gly Val Lys Arg Thr Ala Thr Tyr Leu Pro Glu Val Ala Lys Glu
                230                 235                 240
Gln Asp Trp Asp Gln Ile Gln Thr Ile Asp Ser Leu Leu Arg Lys
                245                 250                 255
Gly Gly Phe Lys Ala Pro Ile Thr Ser Glu Phe Arg Lys Thr Ile
                260                 265                 270
Lys Leu Thr Arg Tyr Arg Ser Glu Lys Val Thr Ile Ser Tyr Ala
                275                 280                 285
Glu Tyr Ile Ala Ser Arg Gln His Cys Phe Gln Asn Gly Thr Leu
                290                 295                 300
His Ala Pro Pro Leu Tyr Asn His Tyr Ser
                305                 310
```

<210> 44
<211> 838
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8227004CD1

<400> 44

```
Met Phe Trp Lys Phe Asp Leu His Ser Ser Ser His Ile Asp Thr
1               5                   10                  15
Leu Leu Glu Arg Glu Asp Val Thr Leu Lys Glu Leu Met Asp Glu
                20                  25                  30
Glu Asp Val Leu Gln Glu Cys Lys Ala Gln Asn Arg Lys Leu Ile
                35                  40                  45
Glu Phe Leu Leu Lys Ala Glu Cys Leu Glu Asp Leu Val Ser Phe
                50                  55                  60
Ile Ile Glu Glu Pro Pro Gln Asp Met Asp Glu Lys Ile Arg Tyr
                65                  70                  75
Lys Tyr Pro Asn Ile Ser Cys Glu Leu Leu Thr Ser Asp Val Ser
                80                  85                  90
Gln Met Asn Asp Arg Leu Gly Glu Asp Glu Ser Leu Leu Met Lys
                95                  100                 105
Leu Tyr Ser Phe Leu Leu Asn Asp Ser Pro Leu Asn Pro Leu Leu
                110                 115                 120
Ala Ser Phe Phe Ser Lys Val Leu Ser Ile Leu Ile Ser Arg Lys
                125                 130                 135
Pro Glu Gln Ile Val Asp Phe Leu Lys Lys Lys His Asp Phe Val
                140                 145                 150
Asp Leu Ile Ile Lys His Ile Gly Thr Ser Ala Ile Met Asp Leu
                155                 160                 165
Leu Leu Arg Leu Leu Thr Cys Ile Glu Pro Pro Gln Pro Arg Gln
                170                 175                 180
Asp Val Leu Asn Trp Leu Asn Glu Glu Lys Ile Ile Gln Arg Leu
                185                 190                 195
Val Glu Ile Val His Pro Ser Gln Glu Glu Asp Arg His Ser Asn
                200                 205                 210
Ala Ser Gln Ser Leu Cys Glu Ile Val Arg Leu Ser Arg Asp Gln
                215                 220                 225
Met Leu Gln Ile Gln Asn Ser Thr Glu Pro Asp Pro Leu Leu Ala
                230                 235                 240
Thr Leu Glu Lys Gln Glu Ile Ile Glu Gln Leu Leu Ser Asn Ile
                245                 250                 255
Phe His Lys Glu Lys Asn Glu Ser Ala Ile Val Ser Ala Ile Gln
                260                 265                 270
Ile Leu Leu Thr Leu Leu Glu Thr Arg Arg Pro Thr Phe Glu Gly
                275                 280                 285
His Ile Glu Ile Cys Pro Pro Gly Met Ser His Ser Ala Cys Ser
                290                 295                 300
Val Asn Lys Ser Val Leu Glu Ala Ile Arg Gly Arg Leu Gly Ser
                305                 310                 315
Phe His Glu Leu Leu Leu Glu Pro Pro Lys Lys Ser Val Met Lys
                320                 325                 330
Thr Thr Trp Gly Val Leu Asp Pro Pro Val Gly Asn Thr Arg Leu
                335                 340                 345
Asn Val Ile Arg Leu Ile Ser Ser Leu Leu Gln Thr Asn Thr Ser
                350                 355                 360
Ser Ile Asn Gly Asp Leu Met Glu Leu Asn Ser Ile Gly Val Ile
                365                 370                 375
Leu Asn Met Phe Phe Lys Tyr Thr Trp Asn Asn Phe Leu His Thr
                380                 385                 390
Gln Val Glu Ile Cys Ile Ala Leu Ile Leu Ala Ser Pro Phe Glu
                395                 400                 405
Asn Thr Glu Asn Ala Thr Ile Thr Asp Gln Asp Ser Thr Gly Asp
                410                 415                 420
Asn Leu Leu Leu Lys His Leu Phe Gln Lys Cys Gln Leu Ile Glu
                425                 430                 435
Arg Ile Leu Glu Ala Trp Glu Met Asn Glu Lys Lys Gln Ala Glu
                440                 445                 450
Gly Gly Arg Arg His Gly Tyr Met Gly His Leu Thr Arg Ile Ala
                455                 460                 465
Asn Cys Ile Val His Ser Thr Asp Lys Gly Pro Asn Ser Ala Leu
```

```
                470                      475                      480
Val Gln Gln Leu Ile Lys Asp Leu Pro Asp Glu Val Arg Glu Arg
                485                      490                      495
Trp Glu Thr Phe Cys Thr Ser Ser Leu Gly Glu Thr Asn Lys Arg
                500                      505                      510
Asn Thr Val Asp Leu Met Gln Gln Met Thr Ser Asn Phe Ile Asp
                515                      520                      525
Gln Phe Gly Phe Asn Asp Glu Lys Phe Ala Asp Gln Asp Asp Ile
                530                      535                      540
Gly Asn Val Ser Phe Asp Arg Val Ser Asp Ile Asn Phe Thr Leu
                545                      550                      555
Asn Thr Asn Glu Ser Gly Asn Ile Ala Leu Phe Glu Ala Cys Cys
                560                      565                      570
Lys Glu Arg Ile Gln Gln Phe Asp Asp Gly Gly Ser Asp Glu Glu
                575                      580                      585
Asp Ile Trp Glu Glu Lys His Ile Ala Phe Thr Pro Glu Ser Gln
                590                      595                      600
Arg Arg Ser Ser Ser Gly Ser Thr Asp Ser Glu Glu Ser Thr Asp
                605                      610                      615
Ser Glu Glu Glu Asp Gly Ala Lys Gln Asp Leu Phe Glu Pro Ser
                620                      625                      630
Ser Ala Asn Thr Glu Asp Lys Met Glu Val Asp Leu Ser Glu Pro
                635                      640                      645
Pro Asn Trp Ser Ala Asn Phe Asp Val Pro Met Glu Thr Thr His
                650                      655                      660
Gly Ala Pro Leu Asp Ser Val Gly Ser Asp Val Trp Ser Thr Glu
                665                      670                      675
Glu Pro Met Pro Thr Lys Glu Thr Gly Trp Ala Ser Phe Ser Glu
                680                      685                      690
Phe Thr Ser Ser Leu Ser Thr Lys Asp Ser Leu Arg Ser Asn Ser
                695                      700                      705
Pro Val Glu Met Glu Thr Ser Thr Glu Pro Met Asp Pro Leu Thr
                710                      715                      720
Pro Ser Ala Ala Ala Leu Ala Val Gln Pro Glu Ala Ala Gly Ser
                725                      730                      735
Val Ala Met Glu Ala Ser Ser Asp Gly Glu Glu Asp Ala Glu Ser
                740                      745                      750
Thr Asp Lys Val Thr Glu Thr Val Met Asn Gly Gly Met Lys Glu
                755                      760                      765
Thr Leu Ser Leu Thr Val Asp Ala Lys Thr Glu Thr Ala Val Phe
                770                      775                      780
Lys Ser Glu Glu Gly Lys Leu Ser Thr Ser Gln Asp Ala Ala Cys
                785                      790                      795
Lys Asp Ala Glu Glu Cys Pro Glu Thr Ala Glu Ala Lys Cys Ala
                800                      805                      810
Ala Pro Arg Pro Pro Ser Ser Ser Pro Glu Gln Arg Thr Gly Gln
                815                      820                      825
Pro Ser Ala Pro Gly Asp Thr Ser Val Asn Gly Pro Val
                830                      835
```

<210> 45
<211> 408
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3044763CD1

<400> 45

```
Met Arg Thr Asp Ser Gly Ala Arg Leu Glu Glu Gly His Leu Arg
 1              5                   10                  15
Pro Pro Arg Ala Leu Pro Pro Val Pro Ser Gln Asp Asp Ile Pro
```

```
                    20                      25                      30
Leu Ser Arg Pro Lys Lys Lys Lys Pro Arg Thr Lys Asn Thr Pro
                35                      40                      45
Ala Ser Ala Ser Leu Glu Gly Leu Ala Gln Thr Ala Gly Arg Arg
                50                      55                      60
Pro Ser Glu Gly Asn Glu Pro Ser Thr Lys Glu Leu Lys Glu His
                65                      70                      75
Pro Glu Ala Pro Val Gln Arg Arg Gln Lys Lys Thr Arg Leu Pro
                80                      85                      90
Leu Glu Leu Glu Thr Ser Ser Thr Gln Lys Lys Ser Ser Ser Ser
                95                      100                     105
Ser Leu Leu Arg Asn Glu Asn Gly Ile Asp Ala Glu Pro Ala Glu
                110                     115                     120
Glu Ala Val Ile Gln Lys Pro Arg Arg Lys Thr Lys Lys Thr Gln
                125                     130                     135
Pro Ala Glu Leu Gln Tyr Ala Asn Glu Leu Gly Val Glu Asp Glu
                140                     145                     150
Asp Ile Ile Thr Asp Glu Gln Thr Thr Val Glu Gln Gln Ser Val
                155                     160                     165
Phe Thr Ala Pro Thr Gly Ile Ser Gln Pro Val Gly Lys Val Phe
                170                     175                     180
Val Glu Lys Ser Arg Arg Phe Gln Ala Ala Asp Arg Ser Glu Leu
                185                     190                     195
Ile Lys Thr Thr Glu Asn Ile Asp Val Ser Met Asp Val Lys Pro
                200                     205                     210
Ser Trp Thr Thr Arg Asp Val Ala Leu Thr Val His Arg Ala Phe
                215                     220                     225
Arg Met Ile Gly Leu Phe Ser His Gly Phe Leu Ala Gly Cys Ala
                230                     235                     240
Val Trp Asn Ile Val Val Ile Tyr Val Leu Ala Gly Asp Gln Leu
                245                     250                     255
Ser Asn Leu Ser Asn Leu Leu Gln Gln Tyr Lys Thr Leu Ala Tyr
                260                     265                     270
Pro Phe Gln Ser Leu Leu Tyr Leu Leu Leu Ala Leu Ser Thr Ile
                275                     280                     285
Ser Ala Phe Asp Arg Ile Asp Phe Ala Lys Ile Ser Val Ala Ile
                290                     295                     300
Arg Asn Phe Leu Ala Leu Asp Pro Thr Ala Leu Ala Ser Phe Leu
                305                     310                     315
Tyr Phe Thr Ala Leu Ile Leu Ser Leu Ser Gln Gln Met Thr Ser
                320                     325                     330
Asp Arg Ile His Leu Tyr Thr Pro Ser Ser Val Asn Gly Ser Leu
                335                     340                     345
Trp Glu Ala Gly Ile Glu Glu Gln Ile Leu Gln Pro Trp Ile Val
                350                     355                     360
Val Asn Leu Val Val Ala Leu Leu Val Gly Leu Ser Trp Leu Phe
                365                     370                     375
Leu Ser Tyr Arg Pro Gly Met Asp Leu Ser Glu Glu Leu Met Phe
                380                     385                     390
Ser Ser Glu Val Glu Glu Tyr Pro Asp Lys Glu Lys Glu Ile Lys
                395                     400                     405
Ala Ser Ser
```

```
<210> 46
<211> 101
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4044519CD1
```

<400> 46
Met Cys Phe Leu Phe Phe Leu Leu Phe Phe Thr Met Val Ala Ser
1 5 10 15
Thr Cys Pro Ser Asp Leu Arg Leu Lys Asp Ser Phe Leu Lys Asn
20 25 30
Met Val Pro Ala Leu Lys Gly Cys Phe Arg Thr Tyr Phe Ile Cys
35 40 45
Phe Leu Leu Ile Leu Ile Phe Gln Leu Asn Pro Ser Ser Ser Leu
50 55 60
Pro Ser Ser Leu Pro Val Tyr Leu Phe Ser Phe Leu Ser Phe Phe
65 70 75
Phe Phe Phe Phe Phe Leu Glu Ala Glu Ser Cys Pro Val Thr Gln
80 85 90
Ala Glu Val Gln Trp Tyr Asp His Ser Ser Leu
95 100

<210> 47
<211> 256
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 71351918CD1

<400> 47
Met Glu Asn Ser Gly Phe Phe Pro Ser Gly Leu Val Val Leu Ser
1 5 10 15
Gly Gly Met Asp Ala Gln Leu Lys Ile Trp Ser Ala Glu Asp Ala
20 25 30
Ser Cys Val Val Thr Phe Lys Gly His Lys Gly Gly Ile Leu Asp
35 40 45
Thr Ala Ile Val Asp Arg Gly Arg Asn Val Val Ser Ala Ser Arg
50 55 60
Asp Gly Thr Ala Arg Leu Trp Asp Cys Gly Arg Ser Gly Cys Leu
65 70 75
Gly Val Leu Ala Asp Cys Gly Ser Ser Ile Asn Gly Val Ala Val
80 85 90
Gly Ala Ala Asp Asn Ser Ile Asn Leu Gly Ser Pro Glu Gln Met
95 100 105
Pro Ser Glu Arg Glu Val Gly Thr Glu Ala Lys Met Leu Leu Leu
110 115 120
Ala Arg Glu Asp Lys Lys Leu Gln Cys Leu Gly Leu Gln Ser Arg
125 130 135
Gln Leu Val Phe Leu Phe Ile Gly Ser Asp Ala Phe Asn Cys Cys
140 145 150
Thr Phe Leu Ser Gly Phe Leu Leu Leu Ala Gly Thr Gln Asp Gly
155 160 165
Asn Ile Tyr Gln Leu Asp Val Arg Ser Pro Arg Ala Pro Val Gln
170 175 180
Val Ile His Arg Ser Gly Ala Pro Val Leu Ser Leu Leu Ser Val
185 190 195
Arg Asp Gly Phe Ile Ala Ser Gln Gly Asp Gly Ser Cys Phe Ile
200 205 210
Val Gln Gln Asp Leu Asp Tyr Val Thr Glu Leu Thr Gly Ala Asp
215 220 225
Cys Asp Pro Val Tyr Lys Val Ala Thr Trp Glu Lys Gln Ile Tyr
230 235 240
Thr Cys Cys Arg Asp Gly Leu Val Arg Arg Tyr Gln Leu Ser Asp
245 250 255
Leu

<210> 48
<211> 104
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8109363CD1

<400> 48
Met Pro Arg Ser Ser Gln His Ser Glu Ser Ser Pro Leu Asp Thr
1 5 10 15
Thr Thr Gln Arg Lys Gly Ala Ser Ser Leu Ala His Gln Val Arg
20 25 30
Val His Thr Leu Glu Thr Leu Leu Asp Trp Pro Glu Leu Pro Gln
35 40 45
Pro Leu Leu Thr Pro Pro Pro Val Ile Asp Thr Ala Ala Gly Ser
50 55 60
Arg Lys Arg Phe Leu Asn Lys Ala Gln Leu Ala Gln Cys Leu Ala
65 70 75
Gln Gln Thr Ile Asn Thr Cys Lys Leu Asn Cys Met Ile Leu Ala
80 85 90
Gln Val Leu Leu Met Trp Leu Thr Ala Thr His Leu His Gly
95 100

<210> 49
<211> 855
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1272746CD1

<400> 49
Met Ser Phe Val Arg Val Asn Arg Cys Gly Pro Arg Val Gly Val
1 5 10 15
Arg Lys Thr Pro Lys Val Lys Lys Lys Lys Thr Ser Val Lys Gln
20 25 30
Glu Trp Asp Asn Thr Val Thr Asp Leu Thr Val His Arg Ala Thr
35 40 45
Pro Glu Asp Leu Val Arg Arg His Glu Ile His Lys Ser Lys Asn
50 55 60
Arg Ala Leu Val His Trp Glu Leu Gln Glu Lys Ala Leu Lys Arg
65 70 75
Lys Trp Arg Lys Gln Lys Pro Glu Thr Leu Asn Leu Glu Lys Arg
80 85 90
Arg Leu Ser Ile Met Lys Glu Ile Leu Ser Asp Gln Tyr Gln Met
95 100 105
Gln Asp Val Leu Glu Lys Ser Asp His Leu Ile Ala Ala Ala Lys
110 115 120
Glu Leu Phe Pro Arg Arg Arg Thr Gly Phe Pro Asn Val Thr Val
125 130 135
Ala Pro Asp Ser Ser Gln Gly Pro Ile Val Val Asn Gln Asp Pro
140 145 150
Ile Thr Gln Ser Ile Phe Asn Glu Ser Val Ile Glu Pro Gln Ala
155 160 165
Leu Asn Asp Val Asp Gly Glu Glu Glu Gly Thr Val Asn Ser Gln
170 175 180
Ser Gly Glu Ser Glu Asn Glu Asn Glu Leu Asp Asn Ser Leu Asn
185 190 195
Ser Gln Ser Asn Thr Asn Thr Asp Arg Phe Leu Gln Gln Leu Thr
200 205 210

Glu Glu Asn Phe Glu Leu Ile Ser Lys Leu Trp Thr Asp Ile Gln
215 220 225
Gln Lys Ile Ala Thr Gln Ser Gln Ile Thr Pro Pro Gly Thr Pro
230 235 240
Ser Ser Ala Leu Ser Ser Gly Glu Gln Arg Ala Ala Leu Asn Ala
245 250 255
Thr Asn Ala Val Lys Arg Leu Gln Thr Arg Leu Gln Pro Glu Glu
260 265 270
Ser Thr Glu Thr Leu Asp Ser Ser Tyr Val Val Gly His Val Leu
275 280 285
Asn Ser Arg Lys Gln Lys Gln Leu Leu Asn Lys Val Lys Arg Lys
290 295 300
Pro Asn Leu His Ala Leu Ser Lys Pro Lys Lys Asn Ile Ser Ser
305 310 315
Gly Ser Thr Thr Ser Ala Asp Leu Pro Asn Arg Thr Asn Ser Asn
320 325 330
Leu Asp Val Leu Lys His Met Ile His Glu Val Glu His Glu Met
335 340 345
Glu Glu Tyr Glu Arg Trp Thr Gly Arg Glu Val Lys Gly Leu Gln
350 355 360
Ser Ser Gln Gly Leu Thr Gly Phe Thr Leu Ser Leu Val Ser Ser
365 370 375
Leu Cys Arg Leu Val Arg Tyr Leu Lys Glu Ser Glu Ile Gln Leu
380 385 390
Arg Lys Glu Val Glu Thr Arg Gln Gln Leu Glu Gln Val Leu Gly
395 400 405
Asp His Arg Glu Leu Ile Asp Ala Leu Thr Ala Glu Ile Leu Arg
410 415 420
Leu Arg Glu Glu Asn Ala Ala Thr Gln Ala Arg Leu Gln Gln Tyr
425 430 435
Met Val Thr Thr Asp Glu Gln Leu Ile Ser Leu Thr His Ala Ile
440 445 450
Lys Asn Cys Pro Val Ile Asn Asn Arg Gln Glu Ile Gln Ala Ser
455 460 465
Glu Ser Gly Ala Thr Gly Arg Arg Val Met Asp Ser Pro Glu Arg
470 475 480
Pro Val Val Asn Ala Asn Val Ser Val Pro Leu Met Phe Arg Glu
485 490 495
Glu Val Ala Glu Phe Pro Gln Glu Glu Leu Pro Val Lys Leu Ser
500 505 510
Gln Val Pro Asp Pro Pro Asp Asn Met Asn Leu Ala Lys Asn Phe
515 520 525
Pro Ala His Ile Phe Glu Pro Ala Val Leu Leu Thr Pro Pro Arg
530 535 540
Gln Lys Ser Asn Leu Lys Phe Ser Pro Leu Gln Asp Val Leu Arg
545 550 555
Arg Thr Val Gln Thr Arg Pro Ala Pro Arg Leu Pro Pro Thr Val
560 565 570
Glu Ile Ile Glu Lys Glu Gln Asn Trp Glu Glu Lys Thr Leu Pro
575 580 585
Ile Asp Thr Asp Ile Gln Asn Ser Ser Glu Glu Asn Arg Leu Phe
590 595 600
Thr Gln Arg Trp Arg Val Ser His Met Gly Glu Asp Leu Glu Asn
605 610 615
Lys Thr Gln Ala Pro Phe Val Asn Leu Ser Gln Pro Leu Cys Asn
620 625 630
Ser His Ser Asn Thr Gln Gln Ser Arg Ser Pro Thr Phe Ser Glu
635 640 645
Glu Leu Pro Val Leu Gly Asp Gly Gln Gln Leu Arg Thr Asn Glu
650 655 660
Ser Leu Ile Gln Arg Lys Asp Ile Met Thr Arg Ile Ala Asp Leu
665 670 675
Thr Leu Gln Asn Ser Ala Ile Lys Ala His Met Asn Asn Ile Ile

```
                    680                 685                 690
Glu Pro Arg Gly Glu Gln Gly Asp Gly Leu Arg Glu Leu Asn Lys
                695                 700                 705
Gln Glu Ser Ala Ser Asp Met Thr Ser Thr Phe Pro Val Ala Gln
                710                 715                 720
Ser Leu Thr Pro Gly Ser Met Glu Glu Arg Ile Ala Glu Leu Asn
                725                 730                 735
Arg Gln Ser Met Glu Ala Arg Gly Lys Leu Leu Gln Leu Ile Glu
                740                 745                 750
Gln Gln Lys Leu Val Gly Leu Asn Leu Ser Pro Pro Met Ser Pro
                755                 760                 765
Val Gln Leu Pro Leu Arg Ala Trp Thr Glu Gly Ala Lys Arg Thr
                770                 775                 780
Ile Glu Val Ser Ile Pro Gly Ala Glu Ala Pro Glu Ser Ser Lys
                785                 790                 795
Cys Ser Thr Val Ser Pro Val Ser Gly Ile Asn Thr Arg Arg Ser
                800                 805                 810
Ser Gly Ala Thr Gly Asn Ser Cys Ser Pro Leu Asn Ala Thr Ser
                815                 820                 825
Gly Ser Gly Arg Phe Thr Pro Leu Asn Pro Arg Ala Lys Ile Glu
                830                 835                 840
Lys Gln Asn Glu Glu Gly Trp Phe Ala Leu Ser Thr His Val Ser
                845                 850                 855


<210> 50
<211> 427
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1839974CD1

<400> 50
Met Tyr Val Thr Met Met Met Thr Asp Gln Ile Pro Leu Glu Leu
 1               5                  10                  15
Pro Pro Leu Leu Asn Gly Glu Val Ala Met Met Pro His Leu Val
                20                  25                  30
Asn Gly Asp Ala Ala Gln Gln Val Ile Leu Val Gln Val Asn Pro
                35                  40                  45
Gly Glu Thr Phe Thr Ile Arg Ala Glu Asp Gly Thr Leu Gln Cys
                50                  55                  60
Ile Gln Gly Pro Ala Glu Val Pro Met Met Ser Pro Asn Gly Ser
                65                  70                  75
Ile Pro Pro Ile His Val Pro Pro Gly Tyr Ile Ser Gln Val Ile
                80                  85                  90
Glu Asp Ser Thr Gly Val Arg Arg Val Val Val Thr Pro Gln Ser
                95                 100                 105
Pro Glu Cys Tyr Pro Pro Ser Tyr Pro Ser Ala Met Ser Pro Thr
               110                 115                 120
His His Leu Pro Pro Tyr Leu Thr His His Pro His Phe Ile His
               125                 130                 135
Asn Ser His Thr Ala Tyr Tyr Pro Pro Val Thr Gly Pro Gly Asp
               140                 145                 150
Met Pro Pro Gln Phe Phe Pro Gln His His Leu Pro His Thr Ile
               155                 160                 165
Tyr Gly Glu Gln Glu Ile Ile Pro Phe Tyr Gly Met Ser Thr Tyr
               170                 175                 180
Ile Thr Arg Glu Asp Gln Tyr Ser Lys Pro Pro His Lys Lys Leu
               185                 190                 195
Lys Asp Arg Gln Ile Asp Arg Gln Asn Arg Leu Asn Ser Pro Pro
               200                 205                 210
```

Ser Ser Ile Tyr Lys Ser Ser Cys Thr Thr Val Tyr Asn Gly Tyr
215 220 225
Gly Lys Gly His Ser Gly Gly Ser Gly Gly Gly Gly Ser Gly Ser
230 235 240
Gly Pro Gly Ile Lys Lys Thr Glu Arg Arg Ala Arg Ser Ser Pro
245 250 255
Lys Ser Asn Asp Ser Asp Leu Gln Glu Tyr Glu Leu Glu Val Lys
260 265 270
Arg Val Gln Asp Ile Leu Ser Gly Ile Glu Lys Pro Gln Val Ser
275 280 285
Asn Ile Gln Ala Arg Ala Val Val Leu Ser Trp Ala Pro Pro Val
290 295 300
Gly Leu Ser Cys Gly Pro His Ser Gly Leu Ser Phe Pro Tyr Ser
305 310 315
Tyr Glu Val Ala Leu Ser Asp Lys Gly Arg Asp Gly Lys Tyr Lys
320 325 330
Ile Ile Tyr Ser Gly Glu Glu Leu Glu Cys Asn Leu Lys Asp Leu
335 340 345
Arg Pro Ala Thr Asp Tyr His Val Arg Val Tyr Ala Met Tyr Asn
350 355 360
Ser Val Lys Gly Ser Cys Ser Glu Pro Val Ser Phe Thr Thr His
365 370 375
Ser Cys Ala Pro Glu Cys Pro Phe Pro Pro Lys Leu Ala His Arg
380 385 390
Ser Lys Ser Ser Leu Thr Leu Gln Trp Lys Ala Pro Ile Asp Asn
395 400 405
Gly Ser Lys Ile Thr Asn Tyr Leu Leu Glu Trp Asp Glu Val Ser
410 415 420
Leu Phe Ser Tyr Ser Pro Ile
425

<210> 51
<211> 800
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1877336CD1

<400> 51
Met Ala Ala Asn Val Gly Asp Gln Arg Ser Thr Asp Trp Ser Ser
1 5 10 15
Gln Tyr Ser Met Val Ala Gly Ala Gly Arg Glu Asn Gly Met Glu
20 25 30
Thr Pro Met His Glu Asn Pro Glu Trp Glu Lys Ala Arg Gln Ala
35 40 45
Leu Ala Ser Ile Ser Lys Ser Gly Ala Ala Gly Gly Ser Ala Lys
50 55 60
Ser Ser Ser Asn Gly Pro Val Ala Ser Ala Gln Tyr Val Ser Gln
65 70 75
Ala Glu Ala Ser Ala Leu Gln Gln Gln Gln Tyr Tyr Gln Trp Tyr
80 85 90
Gln Gln Tyr Asn Tyr Ala Tyr Pro Tyr Ser Tyr Tyr Tyr Pro Met
95 100 105
Ser Met Tyr Gln Ser Tyr Gly Ser Pro Ser Gln Tyr Gly Met Ala
110 115 120
Gly Ser Tyr Gly Ser Ala Thr Pro Gln Gln Pro Ser Ala Pro Gln
125 130 135
His Gln Gly Thr Leu Asn Gln Pro Pro Val Pro Gly Met Asp Glu
140 145 150
Ser Met Ser Tyr Gln Ala Pro Pro Gln Gln Leu Pro Ser Ala Gln
155 160 165

```
Pro Pro Gln Pro Ser Asn Pro Pro His Gly Ala His Thr Leu Asn
                170                 175                 180
Ser Gly Pro Gln Pro Gly Thr Ala Pro Ala Thr Gln His Ser Gln
                185                 190                 195
Ala Gly Pro Ala Thr Gly Gln Ala Tyr Gly Pro His Thr Tyr Thr
                200                 205                 210
Glu Pro Ala Lys Pro Lys Lys Gly Gln Gln Leu Trp Asn Arg Met
                215                 220                 225
Lys Pro Ala Pro Gly Thr Gly Gly Leu Lys Phe Asn Ile Gln Lys
                230                 235                 240
Arg Pro Phe Ala Val Thr Thr Gln Ser Phe Gly Ser Asn Ala Glu
                245                 250                 255
Gly Gln His Ser Gly Phe Gly Pro Gln Pro Asn Pro Glu Lys Val
                260                 265                 270
Gln Asn His Ser Gly Ser Ser Ala Arg Gly Asn Leu Ser Gly Lys
                275                 280                 285
Pro Asp Asp Trp Pro Gln Asp Met Lys Glu Tyr Val Glu Arg Cys
                290                 295                 300
Phe Thr Ala Cys Glu Ser Glu Glu Asp Lys Asp Arg Thr Glu Lys
                305                 310                 315
Leu Leu Lys Glu Val Leu Gln Ala Arg Leu Gln Asp Gly Ser Ala
                320                 325                 330
Tyr Thr Ile Asp Trp Ser Arg Glu Pro Leu Pro Gly Leu Thr Arg
                335                 340                 345
Glu Pro Val Ala Glu Ser Pro Lys Lys Lys Arg Trp Glu Ala Ala
                350                 355                 360
Ser Ser Leu His Pro Pro Arg Gly Ala Gly Ser Ala Thr Arg Gly
                365                 370                 375
Gly Gly Ala Pro Ser Gln Arg Gly Thr Pro Gly Ala Gly Gly Ala
                380                 385                 390
Gly Arg Ala Arg Gly Asn Ser Phe Thr Lys Phe Gly Asn Arg Asn
                395                 400                 405
Val Phe Met Lys Asp Asn Ser Ser Ser Ser Ser Thr Asp Ser Arg
                410                 415                 420
Ser Arg Ser Ser Ser Arg Ser Pro Thr Arg His Phe Arg Arg Ser
                425                 430                 435
Asp Ser His Ser Asp Ser Asp Ser Ser Tyr Ser Gly Asn Glu Cys
                440                 445                 450
His Pro Val Gly Arg Arg Asn Pro Pro Pro Lys Gly Arg Gly Gly
                455                 460                 465
Arg Gly Ala His Met Asp Arg Gly Arg Gly Arg Ala Gln Arg Gly
                470                 475                 480
Lys Arg His Asp Leu Ala Pro Thr Lys Arg Ser Arg Lys Lys Met
                485                 490                 495
Ala Ala Leu Glu Cys Glu Asp Pro Glu Arg Glu Leu Lys Lys Gln
                500                 505                 510
Lys Arg Ala Ala Arg Phe Gln His Gly His Ser Arg Arg Leu Arg
                515                 520                 525
Leu Glu Pro Leu Val Leu Gln Met Ser Ser Leu Glu Ser Ser Gly
                530                 535                 540
Ala Asp Pro Asp Trp Gln Glu Leu Gln Ile Val Gly Thr Cys Pro
                545                 550                 555
Asp Ile Thr Lys His Tyr Leu Arg Leu Thr Cys Ala Pro Asp Pro
                560                 565                 570
Ser Thr Val Arg Pro Val Ala Val Leu Lys Lys Ser Leu Cys Met
                575                 580                 585
Val Lys Cys His Trp Lys Glu Lys Gln Asp Tyr Ala Phe Ala Cys
                590                 595                 600
Glu Gln Met Lys Ser Ile Arg Gln Asp Leu Thr Val Gln Gly Ile
                605                 610                 615
Arg Thr Glu Phe Thr Val Glu Val Tyr Glu Thr His Ala Arg Ile
                620                 625                 630
Ala Leu Glu Lys Gly Asp His Glu Glu Phe Asn Gln Cys Gln Thr
```

```
                635                 640                 645
Gln Leu Lys Ser Leu Tyr Ala Glu Asn Leu Pro Gly Asn Val Gly
                650                 655                 660
Glu Phe Thr Ala Tyr Arg Ile Leu Tyr Tyr Ile Phe Thr Lys Asn
                665                 670                 675
Ser Gly Asp Ile Thr Thr Glu Leu Ala Tyr Leu Thr Arg Glu Leu
                680                 685                 690
Lys Ala Asp Pro Cys Val Ala His Ala Leu Ala Leu Arg Thr Ala
                695                 700                 705
Trp Ala Leu Gly Asn Tyr His Arg Phe Phe Arg Leu Tyr Cys His
                710                 715                 720
Ala Pro Cys Met Ser Gly Tyr Leu Val Asp Lys Phe Ala Asp Arg
                725                 730                 735
Glu Arg Lys Val Ala Leu Lys Ala Met Ile Lys Thr Phe Arg Pro
                740                 745                 750
Ala Leu Pro Val Ser Tyr Leu Gln Ala Glu Leu Ala Phe Glu Gly
                755                 760                 765
Glu Ala Ala Cys Arg Ala Phe Leu Glu Pro Leu Gly Leu Ala Tyr
                770                 775                 780
Thr Gly Pro Asp Asn Ser Ser Ile Asp Cys Arg Leu Ser Leu Ala
                785                 790                 795
Gln Leu Ser Ala Phe
                800

<210> 52
<211> 107
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2321054CD1

<400> 52
Met Ala Gly Gly Tyr Gly Val Met Gly Asp Asp Gly Ser Ile Asp
 1               5                  10                  15
Tyr Thr Val His Glu Ala Trp Asn Glu Ala Thr Asn Val Tyr Leu
                20                  25                  30
Ile Val Ile Leu Val Ser Phe Gly Leu Phe Met Tyr Ala Lys Arg
                35                  40                  45
Asn Lys Arg Arg Ile Met Arg Ile Phe Ser Val Pro Pro Thr Glu
                50                  55                  60
Glu Thr Leu Ser Glu Pro Asn Phe Tyr Asp Thr Ile Ser Lys Ile
                65                  70                  75
Arg Leu Arg Gln Gln Leu Glu Met Tyr Ser Ile Ser Arg Lys Tyr
                80                  85                  90
Asp Tyr Gln Gln Pro Gln Asn Gln Ala Asp Ser Val Gln Leu Ser
                95                 100                 105
Leu Glu


<210> 53
<211> 522
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2796034CD1

<400> 53
Met Thr Pro Gly Lys His Ser Gly Ala Ser Ala Arg Ala Ala Asn
 1               5                  10                  15
```

Ala Gly Ala Trp Gly Tyr Arg Asp Phe Arg Gly Gly Gln Lys Lys
20 25 30
Gly Trp Cys Thr Thr Pro Gln Leu Val Ala Thr Met Pro Val Ser
35 40 45
Pro Ala Gly Ser His Lys Gln Gln Asn Phe Gly Leu Asn Asn Ala
50 55 60
Thr Gln Pro Lys Lys Ser Ile Ser Phe Phe Ala Thr Met Lys Ala
65 70 75
Thr Ser Val Lys Gly Tyr Thr Gly Ala Asn Gln Ser Arg Met Ala
80 85 90
Val Ser Lys Thr Val Leu Ile Pro Pro Glu Leu Lys Thr Val Glu
95 100 105
Lys Pro Asn Pro Asn Ile Lys Thr Thr Gln Val Phe Asp Ile Asn
110 115 120
Gly Thr Asp Val Thr Pro Arg Pro Leu Tyr His Pro Asp Pro Leu
125 130 135
Thr Gly Thr Ala Lys Pro Ser Lys Leu Leu Thr Ser Gln Glu Gly
140 145 150
Ser Leu Gly Ser Glu Phe Ile Ser Ser Tyr Ser Leu Tyr Gln Asn
155 160 165
Thr Ile Asn Pro Ser Thr Leu Gly Gln Phe Thr Arg Ser Val Leu
170 175 180
Gly Ser Ser Thr Val Ser Lys Ser Ser Val Ser Ala Ser Glu Ser
185 190 195
Ile Ala Glu Asp Leu Glu Glu Pro Ser Tyr Lys Arg Glu Arg Leu
200 205 210
Thr Ser Phe Thr Asp Leu Gln Val Ile Arg Ala Ala Pro Glu Lys
215 220 225
Ile Val Thr Lys Glu Asp Leu Glu Lys Asn Ile Glu Ile Ile Leu
230 235 240
Thr Glu Thr Glu Thr Leu Arg Phe Phe Asp Leu Pro Thr Val Met
245 250 255
Val Ser Val Glu Ser Glu Glu Ala Glu Lys Val Thr Gln Arg Asn
260 265 270
Lys Asn Tyr Glu Val Leu Cys Arg Asn Arg Leu Gly Asn Asp Leu
275 280 285
Tyr Val Glu Arg Met Met Gln Thr Phe Asn Gly Ala Pro Lys Asn
290 295 300
Lys Asp Val Gln Cys Asp Lys Ile Ile Met Glu Asp Lys Gly Ile
305 310 315
Met Ser Thr Ala Trp Asp Leu Tyr Asp Ser Tyr Asn Ala Met Glu
320 325 330
Leu Val Ser Leu Ser Val Lys Gln Ser Val Val Glu Ser Ser Ser
335 340 345
Lys Ala Asn Val Leu Pro Lys Asp Gln Asp Gln Arg Leu Pro Gly
350 355 360
Ser Thr Thr Glu Lys Asn Ser Glu Thr Ser Ser Leu Met Asp Ile
365 370 375
Glu Asn Val Ile Leu Ala Lys Ile His Glu Asp Glu Glu Asp His
380 385 390
Ser Asp Ala Ile Leu Lys Ser Asp Lys Phe His Gln Asp Leu Phe
395 400 405
Phe Met Glu Arg Val Leu Met Glu Asn Ile Phe Gln Pro Lys Leu
410 415 420
Ala Ala Tyr Arg Gln Leu Pro Val Leu Lys Glu Pro Glu Pro Glu
425 430 435
Glu Pro Glu Asp Val Leu Glu Ser Ala Lys His Glu Glu Val Glu
440 445 450
Glu Glu Ser Lys Lys Glu Glu Glu Glu Glu Ile His Ala Glu Glu
455 460 465
Ser Thr Ile Pro Ala Asn Leu Glu Arg Leu Trp Ser Phe Ser Cys
470 475 480
Asp Leu Thr Lys Gly Leu Asn Val Ser Ser Leu Ala Trp Asn Lys

485 490 495
Thr Asn Pro Asp Leu Leu Ala Val Gly Tyr Gly His Phe Gly Phe
500 505 510
Lys Glu Gln Lys Glu Asp Trp Leu Ala Ala Gly Gln
515 520

<210> 54
<211> 305
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4413112CD1

<400> 54
Met Gly Gly Thr Leu Ala Trp Thr Leu Leu Leu Pro Leu Leu Leu
1 5 10 15
Arg Glu Ser Asp Ser Leu Glu Pro Ser Cys Thr Val Ser Ser Ala
20 25 30
Asp Val Asp Trp Asn Ala Glu Phe Ser Ala Thr Cys Leu Asn Phe
35 40 45
Ser Gly Leu Ser Leu Ser Leu Pro His Asn Gln Ser Leu Arg Ala
50 55 60
Ser Asn Val Ile Leu Leu Asp Leu Ser Gly Asn Gly Leu Arg Glu
65 70 75
Leu Pro Val Thr Phe Phe Ala His Leu Gln Lys Leu Glu Val Leu
80 85 90
Asn Val Leu Arg Asn Pro Leu Ser Arg Val Asp Gly Ala Leu Ala
95 100 105
Ala Arg Cys Asp Leu Asp Leu Gln Ala Asp Cys Asn Cys Ala Leu
110 115 120
Glu Ser Trp His Asp Ile Arg Arg Asp Asn Cys Ser Gly Gln Lys
125 130 135
Pro Leu Leu Cys Trp Asp Thr Thr Ser Ser Gln His Asn Leu Ser
140 145 150
Ala Phe Leu Glu Val Ser Cys Ala Pro Gly Leu Ala Ser Ala Thr
155 160 165
Ile Gly Ala Val Val Val Ser Gly Cys Leu Leu Leu Gly Leu Ala
170 175 180
Ile Ala Gly Pro Val Leu Ala Trp Arg Leu Trp Arg Cys Arg Val
185 190 195
Ala Arg Ser Arg Glu Leu Asn Lys Pro Trp Ala Ala Gln Asp Gly
200 205 210
Pro Lys Pro Gly Leu Gly Leu Gln Pro Arg Tyr Gly Ser Arg Ser
215 220 225
Ala Pro Lys Pro Gln Val Ala Val Pro Ser Cys Pro Ser Thr Pro
230 235 240
Asp Tyr Glu Asn Met Phe Val Gly Gln Pro Ala Ala Glu His Gln
245 250 255
Trp Asp Glu Gln Gly Ala His Pro Ser Glu Asp Asn Asp Phe Tyr
260 265 270
Ile Asn Tyr Lys Asp Ile Asp Leu Ala Ser Gln Pro Val Tyr Cys
275 280 285
Asn Leu Gln Ser Leu Gly Gln Ala Ser Met Asp Glu Glu Glu Tyr
290 295 300
Val Ile Pro Gly His
305

<210> 55
<211> 329
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7654832CD1

<400> 55
Met Ser Val Leu Glu Glu Asn Arg Pro Phe Ala Gln Gln Leu Ser
1 5 10 15
Asn Val Tyr Phe Thr Ile Leu Ser Leu Phe Cys Phe Lys Leu Phe
20 25 30
Val Lys Ile Ser Leu Ala Ile Leu Ser His Phe Tyr Ile Val Lys
35 40 45
Gly Asn Arg Lys Glu Ala Ala Arg Ile Ala Ala Glu Phe Tyr Gly
50 55 60
Val Thr Gln Gly Gln Gly Ser Trp Ala Asp Arg Ser Pro Leu His
65 70 75
Glu Ala Ala Ser Gln Gly Arg Leu Leu Ala Leu Arg Thr Leu Leu
80 85 90
Ser Gln Gly Tyr Asn Val Asn Ala Val Thr Leu Asp His Val Thr
95 100 105
Pro Leu His Glu Ala Cys Leu Gly Asp His Val Ala Cys Ala Arg
110 115 120
Thr Leu Leu Glu Ala Gly Ala Asn Val Asn Ala Ile Thr Ile Asp
125 130 135
Gly Val Thr Pro Leu Phe Asn Ala Cys Ser Gln Gly Ser Pro Ser
140 145 150
Cys Ala Glu Leu Leu Leu Glu Tyr Gly Ala Lys Ala Gln Leu Glu
155 160 165
Ser Cys Leu Pro Ser Pro Thr His Glu Ala Ala Ser Lys Gly His
170 175 180
His Glu Cys Leu Asp Ile Leu Ile Ser Trp Gly Ile Asp Val Asp
185 190 195
Gln Glu Ile Pro His Leu Gly Thr Pro Leu Tyr Val Ala Cys Met
200 205 210
Ser Gln Gln Phe His Cys Ile Trp Lys Leu Leu Tyr Ala Gly Ala
215 220 225
Asp Val Gln Lys Gly Lys Tyr Trp Asp Thr Pro Leu His Ala Ala
230 235 240
Ala Gln Gln Ser Ser Thr Glu Ile Val Asn Leu Leu Leu Glu Phe
245 250 255
Gly Ala Asp Ile Asn Ala Lys Asn Thr Glu Leu Leu Arg Pro Ile
260 265 270
Asp Val Ala Thr Ser Ser Ser Met Val Glu Arg Ile Leu Leu Gln
275 280 285
His Glu Ala Thr Pro Ser Ser Leu Tyr Gln Leu Cys Arg Leu Cys
290 295 300
Ile Arg Ser Tyr Ile Gly Lys Pro Arg Leu His Leu Ile Pro Gln
305 310 315
Leu Gln Leu Pro Thr Leu Leu Lys Asn Phe Leu Gln Tyr Arg
320 325

<210> 56
<211> 236
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503849CD1

<400> 56
Met Ala Arg Gly Pro Gly Pro Leu Gly Arg Pro Arg Pro Asp Thr
1 5 10 15
Val Ala Met Pro Lys Arg Gly Lys Arg Leu Lys Phe Arg Ala His

```
                20                  25                  30
Asp Ala Cys Ser Gly Arg Val Thr Val Ala Asp Tyr Ala Asn Ser
                35                  40                  45
Asp Pro Ala Val Val Arg Ser Gly Arg Val Lys Lys Ala Val Ala
                50                  55                  60
Asn Ala Val Gln Gln Glu Val Lys Ser Leu Cys Gly Leu Glu Ala
                65                  70                  75
Ser Gln Val Pro Ala Glu Glu Ala Leu Ser Gly Ala Gly Glu Pro
                80                  85                  90
Cys Asp Ile Ile Asp Ser Ser Asp Glu Met Asp Ala Gln Glu Glu
                95                  100                 105
Ser Ile His Glu Arg Thr Val Ser Arg Lys Lys Lys Ser Lys Arg
                110                 115                 120
His Lys Glu Glu Leu Asp Gly Ala Gly Gly Glu Glu Tyr Pro Met
                125                 130                 135
Asp Ile Trp Leu Leu Leu Ala Ser Tyr Ile Arg Pro Glu Asp Ile
                140                 145                 150
Val Asn Phe Ser Leu Ile Cys Lys Asn Ala Trp Thr Val Thr Cys
                155                 160                 165
Thr Ala Ala Phe Trp Thr Arg Leu Tyr Arg Arg His Tyr Thr Leu
                170                 175                 180
Asp Ala Ser Leu Pro Leu Arg Leu Arg Pro Glu Ser Met Glu Lys
                185                 190                 195
Leu Arg Cys Leu Arg Ala Cys Val Ile Arg Ser Leu Tyr His Met
                200                 205                 210
Tyr Glu Pro Phe Ala Ala Arg Ile Ser Lys Asn Pro Ala Ile Pro
                215                 220                 225
Glu Ser Thr Pro Ser Thr Leu Lys Asn Ser Lys
                230                 235
```

<210> 57
<211> 1485
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2867236CB1

<400> 57

```
ctcttcctga gaaacgagca aacctgaaag ctactctctc agcttcagag ggaaaaaatg 60
gttgtagatt tctggacttg ggagcagaca tttcaagaac taatccaaga ggcaaaaccc 120
cgggccacat ggacgctgaa gttggatggc aaccttcagc tagactgcct ggctcaaggg 180
tggaagcaat accaacagag agcatttggc tggttccggt gttcctcctg ccagcgaagt 240
tgggcttccg cccaagtgca gattctgtgc cacacgtact gggagcactg gacatcccag 300
ggtcaggtgc gtatgaggct ctttggccaa aggtgccaga agtgctcctg gtcccaatat 360
gagatgcctg agttctcctc ggatagcacc atgaggattc tgagcaacct ggtgcagcat 420
atactgaaga aatactatgg aaatggcacg aggaagtctc cagaaatgcc agtaatcctg 480
gaagtgtccc tggaaggatc ccatgacaca gccaattgtg aggcatgcac tttgggcatc 540
tgtggacagg gcttaaaaag ctacatgaca aagccgtcca aatccctact cccccaccta 600
aagactggga attcctcacc tggaattggt gctgtgtacc tcgcaaacca agccaagaac 660
cagtcagctg aggcaaaaga ggctaagggg agtgggtatg agaaattagg gcccagtcga 720
gacccagatc cactgaacat ctgtgtcttt attttgctgc ttgtatttat tgtagtcaaa 780
tgctttacat cagaatgatg aaaataggct tgccactttc tcttatttta attccatggt 840
agtcaatgaa ctggctgcca ctttaatata actgaaaatt cattttgaga ccaagcagga 900
tcaagtttgt agaataaaca ctggtttcct agccatcctc tgaaaacagt atgaaacatg 960
accaagtaca taatggattt agtaataaat attgtcgaat tgctaaaaag tcttcaatca 1020
ttcattcact aagtcactca gtgatatcaa tatacttagc tcagaaagtg tgggaggctg 1080
aataatggtg tctcccaaca tatgcatgac ttaatcccca gaacctgtaa acatgttact 1140
ttacatggta gaatggactt tgcggatgta attaaggacc ttgaaatggt tagattattt 1200
catattgtcc gggtggataa gaaccaggat tttgtaacag ggaggcaaca agctcaaaat 1260
cagaaaaaag agatttgtca atggaacaag aggttgaagt gctttgaagt tggaggaaga 1320
ggtcacaggc aaaaaagtac aggcagcctt tagaaaacca aaaggacaaa ggaacagatt 1380
```

ctcccctgga gtctgcagaa ggaaccagcc ctgcctgcac atggctttag cccagtgaca 1440
ctgattttgg acatctgacc ttcagaaata ctagctcata cctcg 1485

<210> 58
<211> 6176
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1294096CB1

<400> 58
caccgctcgg ctccgcgcgg ctctaggagg tggcggcggt ggcggtggcg gcggtggcgg 60
cggcggcggc ggcggggcgc agggctgagc gagcgtccgg gttccggggc tccggggaag 120
gcggttgcag ctcctgagtg cagcgcggct tcctgccact gtcccggccc ggccacctct 180
ctgtcatggc tctggcggac agcacacgtg gattacccaa cgggggcggc ggcgggggcg 240
gcagtggctc ctcgtcgtcc tccgcggagc caccgctctt ccccgacatc gtggagctga 300
acgtgggggg ccaggtgtac gtgacccggc gctgcacggt ggtgtcggtg cccgactcgc 360
tgctctggcg catgttcacg cagcagcagc cgcaggagct ggcccgggac agcaaaggcc 420
gcttctttct ggaccgggac ggcttcctct tccgctacat cctggattac ctgcgggact 480
tgcagctcgt gctgcccgac tacttccccg agcgcagccg gctgcagcgc gaggccgagt 540
acttcgagct gccagagctc gtgcgccgcc tcggggcgcc ccagcagccc ggcccggggc 600
cgccgccctc gcggcgcggg gtgcacaagg agggctcgct gggtgacgag ctgctgccgc 660
ttggctactc ggagcccgaa cagcaggagg gcgcctctgc cggggcgccg tcgcccacgc 720
tggagctggc tagccgcagt ccgtccgggg gcgcggcggg cccgctgctc acgccgtccc 780
agtcgctgga cggcagccgg cgctcgggct acatcaccat cggctaccgc ggctcctaca 840
ccatcgggcg ggacgcgcag gcggacgcca agttccggcg agtggcgcgc atcaccgttt 900
gcggaaagac gtcgctggcc aaggaggtgt ttggggacac cctgaacgaa agccgggacc 960
ccgaccgtcc cccggagcgc tacacctcgc gctattacct caagttcaac ttcctggagc 1020
aggccttcga caagctgtcc gagtcgggct tccacatggt ggcgtgcagc tccacgggca 1080
cctgcgcctt tgccagcagc accgaccaga gcgaggacaa gatctggacc agctacaccg 1140
agtacgtctt ctgcagggag tgagctcccc agaccccctc gccactccag cgcccagtcc 1200
ttctcctgcc cgagagatga ttacagagcc tcttgtccca cctttgtccc ctggctgctg 1260
ccctcccatt ctcccctcc agtagtagct gggtgagacc tgtccgccca ccttccctcc 1320
actacagaac ctgcagccgc aaatcctctg ggctgcttcg tcttctttgg acctcctgaa 1380
ccgagagaac ccagaggaac ccccacccca cccccaccta ccactccatg ctttctctac 1440
tccctgcctc aaaccacccc tcccccagat ggtacttcag tttggatcta ttgggggagt 1500
gtggccacag accgggggat gattgaattg ttcagaacct gattggaccg tgtccaatgt 1560
gcggaagatt tccttgaaat cttctcaagc tcttatgact cactgggggt ttaagagatc 1620
aggattggtt ccactgtctg gggttagtgt tttacaaggt cattacacag tctttttgac 1680
ctcttttgaa ggtagagttt tagaaggctg gatggaagat tctgagcctg gaattaggac 1740
cccatggagg cagttcagta actaaactaa taaagttttg aaaagttaca cgtaaagtag 1800
aagaatctag tgcgtgggac agtaaaggat cctttctcgt acagaataaa aggtctcagc 1860
ctgtagctta aacttataga aagtgatccg cctgcctgca gaggcgccct tttcagctgc 1920
tgctcgccag aagcccttga ttccactggt tgacatggca gcagttactg gcaagaggga 1980
gaaaggacgc tgccgcctaa gagtgcaagg ctgctcaggt ctccaagcgc cgtaggaggt 2040
cacctggcag tgactgtagg gagctgggtc atagtgcacg tcgtgggtat taggaaagcc 2100
tgtattcttt caatgaatgt cagtaggacc ttcctttagc tgtaagactt ggtgggcggg 2160
gtggggtggg gagggaggaa agggtaggaa gggtgggaag ggagaagcag acatagtcat 2220
ttatgatttg aaagttggaa gtttgtacca tctgtttgag tatatgcaca tttaaaaaat 2280
atcatatagt aaatgcaaac atgccaagta ttttataaag attaataaca gacctactct 2340
tacctggcag tttacttaac ttactgtttt gagtcctaaa cttagagttg ttaatgctta 2400
tatataatct aaccaaagag ttacccagta gggttttagt ttttgaactt ttattttctt 2460
gttgattata aatcctgatt ttggaatcta ttgcgcaaaa gaagtttcat tttggttact 2520
tagacctaag atcacttatt aaaaatcctt attttctcca agcccagcaa acgttgactt 2580
ctgggcaaac ctgaaaacct gaaaatgcca ctttcatgca gtttgtttga agttaagtgg 2640
aatcctttca aatgacgagc tgcagagaac tcagcaccaa gggctgccta tctgtagata 2700
gctgtaaaat ggaatatttt taaatgaagg caaataagta cttaaaagtg agctgagcaa 2760
taaaatggtc caataatagg taaatgcaac agaaacagaa ggagacctgg ttgccttatg 2820
cctttactct tacatggaat aaattcccaa tgcatatcct atgtaaacca taagtgaagg 2880
gaaataaacc tcgtcatgct ccatgctgtg aggtgtcctt tggatattct gtgatgacag 2940
agaagcctat tttgttttgt tttcagcatc tttctctgat gtacgttttt aaggattttg 3000

taagagctgt tttcagtgtt taaattagtg ctatttttcc ttgtttttaa aaatgaatct 3060
cgtactgtat cttactatgt ccatacagat gttacaaatc gacagtttta ttcttagact 3120
catgtgatcc aagctgtata taccatatat aaacatttta catgaatcat ttagtttttt 3180
aattcattta ctaatgctat aaaatttcct atattacccc agtaatttgc atcagctggt 3240
ttatatacta aagcaacatg ttttgatgag tttcttacat ccttatcgag gaattgggtt 3300
aggaaaaaat acataattgt aaaactgagt ttgctgtatt atactttttt tcttgagtat 3360
tagttgtatt actaatcata tgttgattaa ctgtctactt aaagtcaagg tacctgtatt 3420
tttaatccac taattttttt ttagttggga aatagatttc aggtctttta ttagactaac 3480
atttttgag aagtaaaatt gacttcatat acaaagcctg taattttagg cgaaatggaa 3540
gcagaaatct aggaagttgt gcttgcttgt atgttgagtt tggtctcaga ctaagtaatg 3600
catcagaatt catctgtttg aagcctgaaa taatttagga ctctgattca ctgaccaaaa 3660
gtcagtgttg cagagatttc tctaccccgt atggtatttt gttagattgt tcaacaggaa 3720
gcacatgatt gagaacatct tgggacagac caaaaccact gacagatggc aaggctcggc 3780
gattctgatt tcccttctca aatctgctca actccaagag tcttgagaaa ctgctaaaat 3840
tttgcctctg tcactcaagt cttacaaatg ttatcttgta aacctttgag gtgaactatt 3900
ccactgtctt gtacataggc atcttattca ctgcaccctg tcacacccag caccccccgc 3960
cccgcacatt atttgaaaga ctgggaattt aatggttagg gacagtaaat ctacttcttt 4020
ttccagggac gactgtcccc tctaaagtta aagtcaatac aagaaaactg tctattttta 4080
gcctaaagta aaggctgtga agaaaattca ttttacattg ggtagacagt aaaaaacaag 4140
taaaataact tgacatgagc acctttagat cccttcccct ccatgggctt tgggccacag 4200
aatgaacctt tgaggcctgt aaagtggatt gtaatttcct ataagctgta atagtggagg 4260
tattgtgggt tcatttgagt aagccctcca aagataccat tcaaataacc tgggagaatg 4320
tcataaatta ttcagataat taacactgca tgaatctgat tcagaggcat gcatttacat 4380
atgttgccct aattaccatt tgatgatcat aaatacaagt gaatgacatt ggacttttag 4440
taacaaactt aatttttaaa aaggtgtaga caatggtggt taaaaaaaaa aaaaaacagg 4500
taccaggttc tgtgtgtttg caccaagtaa ttgacatgtt ttttgtttaa tacatgtgga 4560
ccatgaacag tattcattct cctttttcaa atgatatgct gtagaaaata ttccttgaag 4620
atgtgagatt taaaaatttt tccctttcaa tgttgtttta attgtatttc ttacttggtt 4680
tttttgattg atagcacagt gataaatcat aatactagac aaaattgtct tctctttcaa 4740
accagagcca tatatatgtc tgtatatatg ggacctactg cttctctgag gaaatgcata 4800
atctgttaat atcagacaaa atgagcaatt ggcagtgctc ataatatatt ccaattttta 4860
ttggaatttt cgatggaatg ttatttcaat aaagccatgt aaggtgaaac tttgataact 4920
ttttactctt caagttaggg taaattctga tccaatattc aattcatttg tgtactccca 4980
catgcaaaat gctaaattac aatgcagaca ttaagaaaaa gtattgactg gaggggttga 5040
attccttgag aatttatttt atagtctaaa tcacaaatac tttactcaat ttagttttta 5100
aaatagtaaa ctgaatattt ttgttgtaag cctatcagag tcaatccttc gtttggaatt 5160
gttttcctgt ttttccttac tataaatcat ttaaaaactg aattcatttt cttagatggc 5220
ataagtctgt ctcttgagaa ataagtaaaa tactcctatt ttcagtatct gtagcacctg 5280
aaataggtct ttgtatagcc agaaacaagt tatgttgaag ttagcttttc tttgtcaaca 5340
gttttggaca ataaaaatct gaaagtatta acacttgatt ttctactggg gccccttcaaa 5400
cttggttgga agaaattcaa ccagaatatc tacattagag tataatcatg tgtggtagga 5460
agatggacta gttaatcaag atttgttgtc acttaaattt tttgtgattt ttttccaagc 5520
cagttttttt aaattctaaa tgtgttttga ggtatgggta cattaattgt aatgtaaact 5580
attatacaac tgtttttgcg actttatagg caggtaaatt ttgctattac tattgaatac 5640
aaatgacaat tcatttatga ccactcaaac agcgttagta accatttagt gacaaaggat 5700
taaaacatcc atctggatgt taattttgaa gatgtaaatt atatgttgtt taaatttttc 5760
caggcatctg aaaaccttat ctgctagaca atgtaagatt cacacagagt tatctgggat 5820
tctgattttt taaatagtac atatcattaa accattttct ctaaatgtaa gaagagcaga 5880
aaaaatctta taagattatc agatttttct aatgacacag aaatgtaaga aaaaaatccc 5940
tttatattga aaaaagatgc agtcaaagtc ttttcagaca tgcccaaact ttgagaattt 6000
cttcaaccat ctaatgctat aaagattttt gttcttcctg ttcacaacca gttgtataac 6060
agaaatacta gctactgttt tccttcctgt gtgtgaagta atgaatcatt gattatgtga 6120
cttgttatgt attcaattaa acactaaaga ataaaacatt cactccttta attatt 6176

<210> 59
<211> 1944
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7238537CB1

<400> 59
gggggtgtg ggatggccgc ggagccgggc ggagctggct tgcggctccc ggggccggct 60
ctccggccgg agacatggcc cgggggcccg gcccgctagg caggcctcgc cccgatacgg 120
tcgccatgcc caagagagga aagcgactca agttccgggc ccacgacgcc tgctccggcc 180
gagtgaccgt ggcggattac gccaactcgg atccggcggt cgtgaggtct ggacgagtca 240
agaaagccgt agccaacgct gttcagcagg aagtaaaatc tctttgtggc ttggaagcct 300
ctcaggttcc tgcagaggaa gctctttctg ggctggtga gccctgtgac atcatcgaca 360
gcagtgatga gatggatgcc caggaggaaa gcatccatga gagaactgtc tccagaaaaa 420
agaaaagcaa gagacacaaa gaagaactgg acggggctgg aggagaagag tatcccatgg 480
atatttggct attgctggcc tcctatatcc gtcctgagga cattgtgaat tttccctga 540
tttgtaagaa tgcctggact gtcacttgca ctgctgcctt ttggaccagg ttgtaccgaa 600
ggcactacac gctggatgct tccctgcctt tgcgtctgcg accagagtca atggagaagc 660
tgcgctgtct ccgggcttgt gtgatccgat ctctgtacca tatgtatgag ccatttgctg 720
ctcgaatctc caagaatcca gccattccag aaagcacccc cagcacatta aagaattcca 780
aatgcttact tttctggtgc agaaagattg ttgggaacag acaggaacca atgtgggaat 840
tcaacttcaa gttcaaaaaa cagtccccta ggttaaagag caagtgtaca ggaggattgc 900
agcctcccgt tcagtacgaa gatgttcata ccaatccaga ccaggactgc tgcctactgc 960
aggtcaccac cctcaatttc atctttattc cgattgtcat gggaatgata tttactctgt 1020
ttactatcaa tgtgagcacg gacatgcggc atcatcgagt gagactggtg ttccaagatt 1080
cccctgtcca tggtggtcgg aaactgcgca gtgaacaggg tgtgcaagtc atcctggacc 1140
cagtgcacag cgttcggctc tttgactggt ggcatcctca gtacccattc tccctgagag 1200
cgtagttact gcttcccatc ccttgggggc agcctcgagt gtagtccatt agtaatcaga 1260
ttccagtttg gacagggtgg ctggattgta tatctcgtta gtaatgtaca tgctcttcag 1320
gttctagggc tcctgttagg ggagggagaa atgttgaatc aagagggaaa acaactacta 1380
tgatttataa acatatttta atgtaaaaat ttgcatttaa aaggagtggc cctgttttct 1440
gtgttaaaac cccatttggt gctattgagt ttgttcttta ttcttttatc ccagtgaaaa 1500
ttgttgatct tgctgtaggg aaaaattaaa ctctttgaat ctccaaacaa ggaagtttca 1560
gcattccctt atggatcaga ggaaccttag aggcctgaaa ttgttgcttc cagtttagct 1620
gcccctcaaa ttcaagtgaa tattttccct tctcccttta cccttctcca gaaataaagc 1680
aggtgacagg gttttcagaa tcttaccata ttgacttgtg tatctttttt aaaaaataat 1740
ttgtgatatt tattgcatac attttctttt ggcagttttg gattttaggt attttttttgc 1800
atgaggctga taatggtgat gatgacttta tacccttggt actatactag gaactttaca 1860
tagtatctct acttctcaca gtcttgcaaa gtaggtgata tttattccca ttttatagtt 1920
tgagaaaacg gaattaggtg actg 1944

<210> 60
<211> 1992
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7494391CB1

<220>
<221> unsure
<222> 1962
<223> a, t, c, g, or other

<400> 60
gggcggacag cgatgctcag ctggctgcgg ccgagtcatc gcctagcgct ggcagggccg 60
ctgaccgacc gacggaggcg ccgatggccg attgtccact gcgcagaagg agcagctgct 120
ccgcgccccg ccgcgccgcg ctgaggccga ggtccgcagg gccgcgggga agccgagggc 180
tgccggagaa ccctgcaggt gtcactcggg acgcggaagt gcgcttgccg agaatacagc 240
tacctaccca ggcaatatga agattttatt tgtagaacct gccatttcc ttagtgcatt 300
tgctatgact ttgaccggtc cactgacaac gcaatatgtt tatcggagaa tatgggaaga 360
aactggcaac tacactttttt catctgatag caatatttct gagtgtgaaa aaaacaaaag 420
cagcccaatt tttgcattcc aggaggaagt tcagaaaaaa gtgtcacgtt ttaatctgca 480
gatggacata agtggattaa ttcctggtct agtgtctaca ttcatacttt tgtctattag 540
tgatcactac ggacgaaaat tccctatgat tttgtcttcc gttggtgctc ttgcaaccag 600
cgtttggctc tgtttgcttt gctattttgc ctttccattc cagcttttga ttgcatctac 660
cttcattggt gcattttgtg gcaattatac cacattttgg ggagcttgct ttgcctatat 720
agttgatcag tgtaaagaac acaaacaaaa aacaattcga atagctatca ttgactttct 780

acttggactt gttactggac taacaggact gtcatctggc tattttatta gagagctagg 840
ttttgagtgg tcgtttctaa ttattgctgt gtctcttgct gttaatttga tctatatttt 900
attttttctc ggagatccag tgaaagagtg ttcatctcag aatgttacta tgtcatgtag 960
tgaaggcttc aaaaacctat tttaccgaac ttacatgctt tttaagaatg cttctggtaa 1020
gagacgattt ttgctctgtt tgttactttt tacagtaatc acttattttt ttgtggtaat 1080
tggcattgcc ccaattttta tcctttatga attggattca ccactctgct ggaatgaagt 1140
ttttataggt tatggatcag ctttgggtag tgcctctttt ttgactagtt tcctaggaat 1200
atggcttttt tcttattgta tggaagatat tcatatggcc ttcattggga tttttaccac 1260
gatgacagga atggctatga ccgcgtttgc cagtacaaca ctgatgatgt tttagccag 1320
ggtgccgttc cttttcacta ttgtgccatt ctctgttcta cggtccatgt tgtcaaaagt 1380
ggttcgttcg actgaacaag gtaccctgtt tgcttgtatt gctttcttag aaacacttgg 1440
aggagtcact gcagtttcta cttttaatgg aatttactca gccactgttg cttggtaccc 1500
tggcttcact ttcctgctgt ctgctggtct gttactactt ccagccatca gtctatgtgt 1560
tgtcaagtgt accagctgga atgagggaag ctatgaactt cttatacaag aagaatccag 1620
tgaagatgct tcagacaggt gactgtgatt taaacaaaca aaaaaaatct atgaatgcac 1680
atatcatata ccatgacttc tgaagactat aaatgaattc cacaatcagt gcttcactga 1740
gaaccaattt tacctatctt ttcttctaaa ctgaacagtc agagagacag ctcctggctt 1800
tagcttcttg tggtaccacg cactttgagc actttgtgcg tatcatgcaa tatacttgca 1860
atacacagaa caatttcaaa tacgcctcac ttttagactt agaagagaaa cattaaaact 1920
taagggtgta aggagggatc aagaacttga taaggtcaaa cnatatctct ccttacagcc 1980
aggcaaatca tc 1992

<210> 61
<211> 3906
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6451054CB1

<400> 61
aagcccttgc ctgccaggga actagtatgt cctgtggggg gggagatttt ccctggtgtc 60
tcagggccac agagattgaa caatgggccc gagtcacact gcaagtcagc agcaaaggca 120
gcagtctaaa tgagcccccca aaaagaggag gtacccatgt tctagaaagg tgagccaagg 180
agccccgggg tggcggcagc ggagtaggag gcgaggcaag acctgcggct cggcccggcc 240
acagccgcgg tagtgctagg caagcccacg gagtcacgcc ggcctcagcc agtctgcgaa 300
cctctcgccc ccgcccaccc ctgcatccgc ccggcccctg cccgcccgag tggtccgggg 360
aggcctagtc tgcaccaccc gccgggctcc tggggccgcg cccgcgcggt cctgtcggcg 420
tccgcggtcg cgcccggccc ctggcccgct cgcccgccgg caccatgatg gaggagatcg 480
accggttcca ggtgcccacc gcgcactcgg agatgcagcc gctggaccca gccgccgcct 540
ccatctctga cggagactgt gacgcccggg aggagaagca gcgggagctg gcccggaagg 600
gctccctgaa gaatggcagc atgggtagcc ctgtcaacca gcaacccaag aagaacaatg 660
tcatggcccg aacaaggctg gtcgtcccca ataaaggcta ctcctcactt gaccagagcc 720
ctgatgagaa gccactggta gcccttgaca cggacagcga tgatgacttt gacatgtcta 780
gatactcctc ctccggctac tcctctgctg agcagatcaa ccaagatttg aacatccagc 840
tgctgaagga cggctaccgg ttagatgaga tccccgacga cgaggaccta gacctcatcc 900
cccccaagtc cgtgaacccc acgtgcatgt gctgccaggc cacgtcctcc accgcctgcc 960
acattcagta gcgggcgggg ccactgcggt cgccggggcg gggccgccaa gggccgggta 1020
ggacgggcag aggccgaccc cttcccagct cgtctgcctg cccccggccc cgcgccccta 1080
cagccgccaa cctcgtataa cagtcattgc ttcctcctat ggtaggacgt gtacctctgt 1140
gtgttcatgg agctggagaa accaaaactg ggtctctctc cacccgggag ctgaggaggg 1200
gtgggggccg tttgttcagt tacctggggg aacctcaccc agcaccctgc cccgcttccc 1260
agggctggcc aagagagaag actggacggc agggaggtca acattaacga gtgggggggat 1320
cacaaagcca ggggcttggc ccccacccccag cgaagccatg aactcccaag ccccggcctc 1380
agcctaacca ggaaaggggc tcggagggag aaaggcccag ggcagtgggg acgggtccca 1440
gcacctcgga gcccccttca cacgccccct accggcccat gtttctattt gcatttcccc 1500
ctcctcagat gggaagcctg gcaaagctgc ccaaccgggt agtgcccctc acccatcgga 1560
tcactgcctt cttcctcctc ttgccccatt cacccgcttc cttgcactct gggagtgggg 1620
cagggtaagg tgggccttac agacaggctg aggtttgggt ggtgtgatct gtcctaattg 1680
gcccctcacc aatgcatcta tctgtctgcc accagcccac cccaccctgc ccccaccctc 1740
acccacgctt ctgctcattg gtctcaatcc cagcctggaa gcagggagta aggctggact 1800
ccaatggccc attcccctca ccccagccca agacgtggga ggctgcttcc cagtaagaat 1860

ccaggagcaa gctttggtga atccagtggg ctgagtggct ctatggatgt ggttgtttgg 1920
atgtggtagg gagtgcttag cagaatgtgt gtgtatctct gaggctgtca gggggtgggt 1980
gggtgtgtgt ctctgttgga cctgtctctc ctcagtctat tcctgtcata gatgtgggtg 2040
tccctggtag ggaagtgcct gtgtggtctg tttacagtgc agaaggcggc tcagcagtgc 2100
aatgggctga cagccttgga ccgggagtca gaagcctggg ttctggttgt agctctggcc 2160
cctccctgtg atcttcaaca agccacttgc tttctctcgt cttcagtttc ctcaactaga 2220
aaatgaacag cagctgaaat tatgttgaag atccttttat acctgaaaga gtttcttcat 2280
agaaatttct gtggtgggga gggggcaatc tctgtgtgtc tcagtatacc taaaggttgg 2340
ggtggggcgg agttgtgtgt ctctgtagaa aggcgaaggg gcagttccta tcaggactgt 2400
gtatgtctga gcacatgtgg ctctgtttgg gattacgtgt ttgtctgtga atgtgtgtgt 2460
gtgttggagg gttgtctatt gtgtgtggct gtatagggtg tctgtagatc aagatgtgta 2520
tacagctgct tctgctattg ctggtttggg ggaggggact ggaaagctga gactgaaaat 2580
caagagggaa gaggtgagga aggcccaccc ctggggctgg gcagtgtaga cctgccaccg 2640
tgtgaagaaa aatgaaagca cagatagcct aagactgcac aggcctgtgc ccacatacct 2700
tatagcacag gcccaggtgg atcatctcat gctaacaccc cacatgtgct ggtggatcca 2760
ctcccatgga tccactcaca tgctgcaggc atgcagtaga ctcacccaca gcccactctc 2820
tggacaccaa tcctgggagg tgggagctgg aatggggagc aatgagagga gggagtgaga 2880
ctgttggctt aggaagggtg gtacctctcc ctgcatcata cccaggaact tccagctcca 2940
aaatattggg gtacagctca ccacccccca ctccgccaac ccattcctac ctcctctagt 3000
gcagggaccc aagttagatg atgagaaaat gccaaacaaa gggaggcaag gacagagaag 3060
gtacctgcag gacccacaag gcccatactg gcccccatct gaacccacct ctcccaccca 3120
gatcaacatc cccccaaccc ctgtgccccc atcatctgga ccccagagcc agccctccca 3180
ctaccaccac caccaccccg gggctccttg gctcaccttt ctagaacatg ggtacctcct 3240
ctttttgggg gctcatttag actgctgcct ttgctgctga cttgcagtgt gactcgggcc 3300
cattgttcaa tctctgtggc cctgagacac cagggaaaat ctcccccccc cacaggacat 3360
actagttccc tggcaggcaa gggcttccaa ctgaggcagt gcatgtgtgg cagagagagg 3420
caggaagctg gcagtggcag cttctgtgtc tagggagggg tgtggctccc tccttccctg 3480
tctgggaggt tggagggaag aatctaggcc ttagcttgcc ctcctgccac ccttcccctt 3540
gtagatactg ccttaacact ccctcctctc tcagctgtgg ctgccaccca gccaggtttc 3600
tccgtgctca ctaatttatt tccaggaagg tgtgtggaag acatgagccg tgtataatat 3660
ttttttaac atttcattg cagtattgac catcatcctt ggttgtgtat cgtgtaacac 3720
aaataatgat attaaaagca tcaaacaaaa aaaaaataaa tggtgctgac gcctctggct 3780
tgcgccagga cttggcgtgg gcaccgggcg cccccatccc agtgtctgtg tgcgtccagc 3840
tgtgttgcac aggcctgggc tccccactga gtgccaaggg tcccctgagc atgcttttct 3900
gaagaa 3906

<210> 62
<211> 3236
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7494592CB1

<220>
<221> unsure
<222> 3215
<223> a, t, c, g, or other

<400> 62
gccggacggc tgacagagcg gctgacggag ccgggctcac caggtcgctg ccgcgaggga 60
gttgctgtgc tggggcctgg gtggcggctg gaggcctgag ttgggctcgc ggcgggggtc 120
gcgcaggggg cctgggtggc ggaatgatgg aggaggagga actggagttc gtggaggagc 180
tggaagccgt gctgcagctc acgcccgagg tgcagctggc catcgagcag gtgtttccaa 240
gccaggaccc tctagatcga gcagatttca atgctgttga gtatatcaat accctgttcc 300
caaccgagca atctctggcg aacatagacg aagtcgtgaa caaaattagg ctgaaaataa 360
ggagactgga tgacaatatt cgaactgttg taagaggtca gacgaacgtg gggcaggatg 420
gacggcaagc gcttgaagag gctcagaaag ctatccaaca actctttggc aaaatcaaag 480
atatcaaaga caaagctgaa aaatcagagc aaatggtgaa agaaatcacc cgtgatatta 540
agcaattaga tcacgccaaa cgccacctga ccacctcaat caccacactg aaccacctgc 600
acatgctggc aggaggtgtc gactccctcg aagccatgac caggcgaaga caatacggag 660
aagttgctaa tctccttcag ggtgtgatga atgtcctgga gcacttccac aagtatatgg 720

ggattccgca gatccggcag ctttccgaaa gagtgaaggc tgcacagact gagttaggac 780
agcaaatcct ggcagatttt gaagaagcgt ttccttccca gggcaccaag agaccaggag 840
gacccagcaa tgttctacga gatgcatgtc tggttgctaa tattctagat cccaggatca 900
aacaggaaat catcaaaaag tttattaaac agcatctgtc agagtatctg gtactttttc 960
aagaaaacca agatgttgcc tggctggaca aaatcgacag acgctatgcc tggataaaac 1020
gccagcttgt ggactatgag gagaaatacg gccgcatgtt tccacgtgag tggtgcatgg 1080
ctgagaggat tgcggtggaa ttttgccatg tgacaagggc agaacttgcc aagattatgc 1140
gtaccagagc gaaggaaatt gaagtgaaat tgcttctttt tgctattcaa agaacaacta 1200
actttgaggg gtttcttgca aaacgcttct ccggctgcac cctgaccgat gggaccctga 1260
aaaagcttga gtctccaccc ccatctacca atcccttcct ggaagatgag ccaacaccag 1320
agatggagga actggcaacg gagaaaggag atttagatca accaaagaag cctaaagccc 1380
cagacaatcc atttcatggc attgtttcca agtgttttga gcctcatctc tacgtgtata 1440
tcgaatccca agacaagaac ctcggagagc tgatagatcg gtttgtggct gatttcaaag 1500
cccaggggcc acctaagccc aacactgatg aaggggtgc cgtgctcccc agctgcgccg 1560
acctctttgt ctactacaag aagtgcatgg tgcaatgctc tcagctcagt actggggagc 1620
ccatgatcgc cctgaccacc attttccaga agtacctccg agaatacgcc tggaaaatcc 1680
tctctggcaa cctgcccaaa accacaacca gcagtggagg actgactatc agcagcctcc 1740
tcaaggaaaa ggagggctca gaagtagcca agttcactct ggaggagctc tgcctcatct 1800
gtaacatcct gagcacggca gagtactgtc tggccaccac ccagcagcta gaagaaaaac 1860
tcaaagaaaa agtggatgta agtctgattg aacgaatcaa tctgactgga gagatggaca 1920
cgttcagcac cgtcatctcc agcagtattc agctgctggt tcaggatctg gatgctgcct 1980
gtgatcctgc cctgactgcc atgagcaaga tgcagtggca gaacgtggag cacgttggtg 2040
accagagccc ctacgtcacc tctgtcattc tgcacatcaa gcagaacgtc cccatcatcc 2100
gtgacaacct ggcttccaca cgcaagtact tcactcagtt ctgcgttaaa tttgcaaact 2160
ccttcattcc caaattcatc acccacctct tcaagtgcaa gccaattagc atggtgggag 2220
cagaacagct gctgctggac acccactcgc tgaagatggt cctgctcgat ctcccctcca 2280
tcagctcgca ggtggtgagg aaggcacccg ccagctacac caagatcgtt gtcaaaggca 2340
tgacccgggc tgagatgatc ctcaaggtag tgatggcccc tcatgaaccg ttggtggtgt 2400
ttgttgacaa ctacatcaaa cttctcacag actgcaacac agaaaccttt cagaagatac 2460
tggacatgaa ggggctgaag aggagtgagc agagcagcat gctggaactc ctgcgccagc 2520
ggctccccgc accgccctcg gggcagaaa gctccggctc actgtccctg acggcgccaa 2580
caccagagca agagtcgtca cgcatccgca agctcgagaa actcattaaa aagagactgt 2640
agcagcagca aggggccctt tgctcctggc tggagaccct cagcgcccgt tccccagaag 2700
cccccaacct ctcctgtgct ccccggcact ctcacatcgt cggtcttcaa acttcctggg 2760
acatgtgggt tgttactgag tctctcccat gccctgtctt acttcctgcc ctaatcggag 2820
acgctgagta agggctgggc tctaagaggg cgatttaggt gatctctggt tcgtgaagca 2880
gaggcagcag tggaggagtt tgagaccagt ctgggcaacc tggcaaaacc aaaaatacaa 2940
aaaattggct gggtgtggtg gctcatgcct gtaatcacag cactttggga ggctgaggca 3000
ggtggatcac ctgaggtcag gaattcgaga ccagcttgcc caacatggcg aaaccctgtc 3060
tcaactaaaa ataaaaaaat ttacccgggg tcaatgcgaa tgcctgtatt cccagtactg 3120
ggaagctaag gaagggattc acttgaaccc gggagttgaa ggttaattaa ccaagtttca 3180
ccaatgaatt caacctggtg aaaagcaaac ctctntcaaa agtttacggt taggga 3236

<210> 63
<211> 1906
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5202657CB1

<400> 63
cggaagccgg ggccggggct gcggggcgag ttgtcggccc tgggccggga gctggagtcc 60
cagactcata ggtcccggcc cagccccga agagccgcct cagccggggg gagttgctcg 120
gactcaaacg tccagtcctc gtgcgaccgc gctgggtcgg aagtgagcag gctgaggcca 180
ccatggagca gtgtgcgtgc gtggagagag agctggacaa ggtcctgcag aagttcctga 240
cctacgggca gcactgtgag cggagcctgg aggagctgct gcactacgtg ggccagctgc 300
gggctgagct ggccagcgca gccctccagg ggacccctct ctcagccacc ctctctctgg 360
tgatgtcaca gtgctgccgg aagatcaaag atacggtgca gaaactggct tcggaccata 420
aggacattca cagcagtgta tcccgagtgg gcaaagccat tgacaggaac ttcgactctg 480
agatctgtgg tgttgtgtca gatgcggtgt gggacgcgcg ggaacagcag cagcagatcc 540
tgcagatggc catcgtggaa cacctgtatc agcagggcat gctcagcgtg gccgaggagc 600

tgtgccagga atcaacgctg aatgtggact tggatttcaa gcagcctttc ctagagttga 660
atcgaatcct ggaagccctg cacgaacaag acctgggtcc tgcgttggaa tgggccgtct 720
cccacaggca gcgcctgctg gaactcaaca gctccctgga gttcaagctg caccgactgc 780
acttcatccg cctcttggca ggaggccccg cgaagcagct ggaggccctc agctatgctc 840
ggcacttcca gccctttgct cggctgcacc agcgggagat ccaggtgatg atgggcagcc 900
tggtgtacct gcggctgggc ttggagaagt caccctactg ccacctgctg gacagcagcc 960
actgggcaga gatctgtgag acctttaccc gggacgcctg ttccctgctg gggctttctg 1020
tggagtcccc ccttagcgtc agctttgcct ctggctgtgt ggcgctgcct gtgttgatga 1080
acatcaaggc tgtgattgag cagcggcagt gcactggggt ctggaatcac aaggacgagt 1140
taccgattga gattgaacta ggcatgaagt gctggtacca ctccgtgttc gcttgcccca 1200
tcctccgcca gcagacgtca gattccaacc ctcccatcaa gctcatctgt ggccatgtta 1260
tctcccgaga tgcactcaat aagctcatta atggaggaaa gctgaagtgt ccctactgtc 1320
ccatggagca gaacccggca gatgggaaac gcatcatatt ctgattccta cctggaagga 1380
atttgttga aaggggtttt cacctgtgag ccttggtctg tctcggtagg gtggtcaact 1440
tcagtggact gtggttggtt tcagagcgcc tggctgagga gttccactga ggggagcact 1500
ggagcagccc tttggcagag gctgaggagg gagatggacc agcccacgcc tggcacctgg 1560
ctccatggca taaggaaagg gagatgctgg cctctgtgct cctgctgtct tttcctgttt 1620
ctgtttgcgt ttgacttagt agcaaccgac agagtggcaa gggatttggt cttcagcagt 1680
agacatcctt ccacccctgc cctcagccaa gtctcttgct gccatgccaa tgctatgtcc 1740
acccttgccc ctcggcccaa gagtgtccag cggtggccca cctcttcctc ccactacagc 1800
ctcaacagta tgtaccatct cccactgtaa atagtcccag ttagaacgga atgccgttgt 1860
tttataactt tgaacaaatg tatttactgc ccttctcaaa aaaaaa 1906

<210> 64
<211> 1347
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2013529CB1

<400> 64
gcccgggcgc aggcggcgga tggagcggaa cggctagggg tcttgagaag caatggccac 60
agaagcccct gtgaatatag caccacctga gtgtagcact gttgtcagca cagcagttga 120
cagcctcatt tggcagccaa actcactaaa tatgcacatg ataaggccca agtccgccaa 180
gggacggaca agaccgagtc tgcagaaatc ccagggcgtg gaggtgtgcg ctcatcatat 240
accatctccg cctccagcca ttccctatga gttgccaagc agccaaaaac caggagcctg 300
tgcacccaaa tctccaaacc agggagcttc tgatgagatc cctgagctgc agcagcaagt 360
acccactggg gcttcctctt ctctcaataa gtatccagtc cttccttcca tcaacagaaa 420
gaacctggag gaggaggctg tggaaaccgt tgccaaaaag gccagctcac tgcaactgag 480
cagtatccgg gctctttacc aagacgagac gggcaccatg aagacaagtg aagaagattc 540
cagagctcga gcttgtgccg tggagaggaa attcatcgtc cgaaccaaga aacagggctc 600
ttccagggct ggaaatctgg aggaaccatc ggaccaagaa ccaaggttgc tgcttgctgt 660
tagatcacca acaggccaaa ggtttgtacg ccatttccgg ccaacagatg atttgcaaac 720
cattgttgct gtggccgaac agaaaaacaa aacctcctac cgacactgca gcattgaaac 780
aatggaggtg cccaggaggc gattttctga cctcaccaaa tctctgcaag agtgcagaat 840
cccccacaag tctgtgctgg gcatctcact ggaagatggg gaagggtggc cctgagtcca 900
cagccaccca gctgaggtcc tgggtctctg agcaaaggag catgcttggg cgttgtggcc 960
tcttaggcag cctgtttcaa gtgccatgtg gacctggtgc agctgggaag cttgggactc 1020
tcgtctgcac tgcgtgtcct ctgaagcagt gaagtctgtg cctatgccga gcgcgctaag 1080
aagtctccct tccagctgtt ccattctctc caccaccagc gtaactggca agttaccaag 1140
gttgttcctg aaacagcagt gatcatgact tctcctttcc agagttttgg gtccttctga 1200
attaatggtc cttttcgaac accggcttgc ctttacagtg aactgtgatt ctctcgaagc 1260
caatgctttc ctgtctgtat ttgatgcagg attaaacact tcccagagag gattctagtc 1320
tggtaaataa ccaccaaaaa aaaaaaa 1347

<210> 65
<211> 1854
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 3841351CB1

<400> 65
gcagagcgag tcgggaaacg attttaaact gaagaggcgg cggagggccg aattcccttt 60
tctcaacggc ttgatttcag agctgggctg gtctctgaca ggctcagctg gagagggacg 120
ggttgggacg cactgtcctt ttgcccttcc ccctccgcga gcagaagctg actccgcagg 180
agcgagggtc gcagagctgg gggatttcag tctccacata gttttggagc cggacttttg 240
aagaatgatt cgtgaatccg gaatgggtga cagcgtcatc acggcatttt attgacagac 300
catggattct tacagtgcac cagagtcaac tcctagtgca tcctcaagac ctgaagatta 360
ctttataggt gccactcctc tgcagaaacg attagaatcg gtcaggaagc agagttcatt 420
tatcctgact ccacctcgaa ggaaaattcc ccagtgttcg cagttgcagg aagatgttga 480
ccctcaaaag gttgcattcc ttctgcataa acagtggact ttatatagtt taactccctt 540
atataaattc tcctatagta atctcaaaga gtattctaga cttctcaatg cttttattgt 600
tgctgaaaag caaaaaggac ttgctgtgga agtgggagaa gacttcaaca tcaaagtgat 660
tttttctact ctcctaggaa tgaaaggaac acaaagggac ccggaagcat ttcttgtcca 720
gattgtgtca aaatctcaat tgccatctga gaatagagaa ggtaaagtgc tgtggactgg 780
ctggttctgc tgtgtatttg gagacagtct tctggagact gtttcagaag atttcacctg 840
tctgccctta ttccttgcaa atggagcaga gtctaacaca gcaataattg gaacttggtt 900
tcagaaaacc tttgactgtt atttcagtcc tttagcaatc aatgcattta atctttcctg 960
gatggctgcc atgtggactg catgcaaaat ggaccattat gtggctacta ctgaatttct 1020
ttggtctgta ccctgtagcc ctcaaagtct ggacatttct ttcgcaatac atccagagga 1080
tgcaaaagct ctatggaca gtgtccacaa aacacctggg gaggttaccc aggaagaagt 1140
tgacctattc atggattgcc tttattcaca tttccataga catttcaaaa ttcatttatc 1200
agccacaaga ttagttcgtg tttcaacatc tgtagcttca gcacatactg atggaaaaat 1260
aaagattctg tgtcataaat accttattgg agtgttagca tatttgacag aactggcaat 1320
ttttcaaatt gagtgaagcc ttatgtggac tataagttat agattatata ctcttattga 1380
taacttgcct aattgctatg ctgaaagaga ctgcaggaga aataggcatc tatctctgca 1440
tctgttttcc ccaccatgcc tttggagttg ccaagatgga agccaagaag gatctagaag 1500
aacaaagaat atggtagtag atgagccaca gccaggtgcc catgtactaa tcatgataac 1560
ctgacatgcc attctcaaaa tgctgagttg ttaatttctt gtcatcttta aatatatata 1620
tataggctgg gcttggtggc tcacacctgt aattccagca ctttgggagg ctgaggtggg 1680
tggatcattt gaggccagga attcaagacc agcctggcca acatggtgaa accccttctc 1740
tactgaaaat acaataatta gctgggcgtg gtggcacatg cctatgatcc cagctactgg 1800
ggaggctgag gcaggagaat cgttttaacc cagaagacag gctgcagtga gcca 1854

<210> 66
<211> 1327
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 152116CB1

<400> 66
tgcctcctct ggctgctgcc tccgcagctc cctcctccta ccccacctcc tccatctggg 60
gagcgtctgc gggggcctga ggggcggcgg cggcggcggc ggctgcgata tggagccagg 120
ggccggcggc cgaaatactg ctcgtgcgca gagggccggg tccccgaaca ctccgccgcc 180
ccgggagcag gagcggaaac tggagcagga gaagctctcc ggtgtggtga agagcgtcca 240
ccggcggctc cgcaagaagt accgagaagt gggagatttt gataagatct ggcgagaaca 300
ctgtgaggat gaagaaactc tttgtgaata tgccgttgca atgaaaaatt tggcagataa 360
tcattgggca aaaacttgtg agggcgaagg tcgtattgaa tggtgttgta gtgtatgcag 420
agaatatttc caaaatggtg ggaagagaaa agcacttgag aaagatgaaa aaagagctgt 480
acttgccact aagaccactc cagccttaaa tatgcatgag tcttctcaac ttgaaggtca 540
tttaaccaac ctcagcttta caaaccctga atttataact gagttgctac aagcctcagg 600
aaaaatcaga ttacttgatg ttggcagctg ctttaaccca tttctgaagt ttgaagaatt 660
tctaactgtt ggcatagata ttgtacctgc tgtagagagt gtctataaat gtgatttcct 720
gaacttacag cttcagcaac cactccagct tgcacaggat gctatagatg ctttttttgaa 780
gcagctgaaa aaccctattg attctcttcc tggagagctt ttccatgtgg ttgttttctc 840
tctccttctt tcttattttc catctcctta ccagcgatgg atttgctgca agaaagccca 900
tgaactgtta gtgttaaatg gtttattact aatcatcaca cctgattcct cccatcagaa 960
ccgtcatgct atgatgatga aaagctggaa gattgctata gagtccctgg gctttaaacg 1020

cttcaagtac tcaaaatttt cacatatgca tctgatggca tttaggaaaa tctctctaaa 1080
aaccacaagt gacttggtta gtaggaacta cccaggaatg ttatatattc ctcaagattt 1140
caacagtata gaagatgagg aatattctaa cccttcctgc tatgttcgat cagatataga 1200
agatgaacaa ctagcatatg gtttcacaga actccctgat gcgccatatg actcagattc 1260
tggagaaagt caagccagct ctattccttt ctatgagcta gaagacccca tattactttt 1320
aagttaa 1327

<210> 67
<211> 627
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2381031CB1

<400> 67
atccgggcct gagagtgcag gcttgaggga agcatggagg tccatggcaa gcccaaggct 60
agcccgagtt gttcgtcgcc cacccgggat tcctcaggag tcccagtgtc caaggagctg 120
ctgacggcgg gaagcgacgg ccgcggaggt atatgggaca ggttgctcat caactcccaa 180
cctaagtcca gaaagacctc cactcttcaa acagttcgga tagagaggag tcccttattg 240
gaccaggtac agacatttct cccacagatg gcacgggcaa atgaaaagct aagaaaagaa 300
atggcagctg caccacctgg tcgtttcaat attgaaaaca ttgatgggcc tcatagtaaa 360
gttatacaaa tggatgtggc tttgtttgag atgaatcagt cggattcaaa agaagtggac 420
agttcagaag agagttcaca agacagttca gagaacagtt cagaatcaga agacgaagat 480
gacagcatcc catctgaagt caccatagat aacattaagc ttcccaattc tgaaggtgga 540
aaaggcaaga ttgaagtttt ggacagtcca gcaagtaaaa aaaagaaata gtcaaataaa 600
ttatctgaaa agaaaaaaaa aaaaaaa 627

<210> 68
<211> 2564
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2511371CB1

<400> 68
gcacgctgac tttgctctca ctgcccggcc agacgcactt cccgccttcg cctaggtctg 60
caccgccacc gccgccctct cccgcgcgtc cctcacccgg tccgccactg ccgcgcctca 120
gaatccgaag gctcggttcc agcgcagcca tccacgccgc cgccacgacc gactccggag 180
ccacgcgctg caaccgccac caacgccttg ccgccgccgc tcgcccgccg accaccatgg 240
cccagcagca aacaggtagc aggaaacgga aagcgcccgc ggtcgaggcg gacgccgaga 300
gctcgccgtc gcagggcttg gcggcagcgg acggtgaggg gccgctgcta ctcaagaggc 360
agaggcggcc ggcgacgtat cgctcgatgg cgcactatct gaaggttcgg gaggtaggcg 420
ggtggggccc cgccaggctc cagggcttcg atggcgagct gcggggctac gcggtacaga 480
ggctgcccga gctgctgacg gagcgccaac tggagctggg cacggtcaac aaggtgttcg 540
cgtcacagtg gctgaactcc aggcaggtgg tgtgcggcac caagtgtaac acgcttttcg 600
tggtggacgt ggagtcaggc cacatcgcgc gcattcccct cttgcgggac agtgaggcca 660
ggctggccca ggaccaacag ggctgcggca tccatgccat cgagctgaat ccctccaaga 720
cgcttctggc caccggcggc gaaaacccca acagcctggc catctaccag ctgccctccc 780
tggatcccct gtgcctgggc gaccgccatg gccacaagga ctggatcttc gccgtcgcct 840
ggctgagtga caccgtagcc gtgagcggct cccgcgacgg cactgtggcg ctgtggcgga 900
tggacccgga caagttcgat gacactgttg cctggcatag cgaggtgggt ctcccccgtat 960
atgcccacat ccgtccgagg gatgtggagg ccatccccag ggccatcatc aaccccagta 1020
accgcaaggt gcgggccctg gcctgcggcg gcaagaacca ggaactggga gcggtgtcct 1080
tggacggcta cttccacctg tggaaagccg ggagcgcact atccaggctg ctgtccatca 1140
ggctgcccta cttccgggat aatgtgtgcc tgacctactg tgatgatatg tctgtgtacg 1200
ccgtgggctc ccattcccac gtctctttcc tggatctgcg ccaggaccag cagaacatcc 1260
ggcccctgtg ttctcgagag ggtggcacag gcgtgcggtc gctgagcttc taccgccaca 1320
tcatcactgt gggcaccggt cagggctccc tgctcttcta tgacgtccgg gcccagaaat 1380
tcctggagga aagagcctcc gccaccctgg agtcctcttc gggacctgca aggaggaagc 1440

tcaggcttgc ctgtggcaga ggctggctca accacaatga tttctgggtg aattactttg 1500
gtggcatgga agtgtttccc aatgcgctct acacccactg ctacaactgg cctgagatga 1560
agctctttgt ggctgggggg cctctccctg caggcctcca tgggaactat gcaggcctct 1620
ggagctaagg atgaccgcct tgtactccag gtggaaccag cccagcttaa tttgtactcc 1680
tccaagcaga gagagtgagt ggtcctttgg tcaatgaaaa cttggcttta agacttttg 1740
actcctttcc aacaccatct tttggaattt tcatacaaag aaatttggtt aatcgttaat 1800
ctttagttgc attctctctt tgagtgattc atacattctg tttctttaag aagatgcagt 1860
catggtcttt cagttatgta tatttgtttt ccttgttttt cctcgttacc aatattacaa 1920
cagcattgtt ttcagttatt tttcattgca gatacgaaga gaaaattttc tagtaagcaa 1980
ggtgttccta catcttttct ggagagctta ccatagtttt atagcattgt tcaattggaa 2040
aagtacattc caagttctag attacctacg tttggaaaca cttaatttac aaattgtgat 2100
cttagtgtat atagttttta cgttgttggt gctaaagtgc ttttttgaag tttattctct 2160
ttcaaaataa gttatcatag aatcttgtgc ttgtatcaaa ctcctaattg gtcggaacac 2220
tttaggagtg caggggttct gcttcattgc attttcagaa ttaacctgaa agttgaaaca 2280
accatttttt tttatttcaa aacttattta agttgttatc acctgtcctt aaatagataa 2340
aagaatgtgt atttttttcac ttgttgcctt agtactatca agacttgtga atatatttga 2400
atgtatgttg tttgaaaaac ttgcacaggg cattggagcc atcatgaaaa gcaggaaatg 2460
taaacgtgga atgtatggca ggctgggcag aatgtatact gatccagggt atagactaga 2520
tatttttgct gttgctgttt ttattataaa atgttatatt tgag 2564

<210> 69
<211> 4134
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8068623CB1

<400> 69
tgctggtgct gtgccgtcgg ccctgccggc ggcgcagctc gggctgcgtt ccttctctgc 60
gtgcgcttgg cactgcgtcg tcgtctttct ggtttgtcct ctgtggttct ctccccgctc 120
ccttgtctgt gtcttcggct gtgctggtcg ggttcgttgt cgttgtgtgt tcctctgtcg 180
tgtcttgctc tgcgtttgct gcttgcgttg tctgtgctcg tcgtgtgctc gctgtgttgt 240
ggtgtgtttg tgctctggtt ggcgtttgct tgtcgtcctt gggcgtcggt gcttggttgt 300
gtcggttttc tcgtatgact gtcacgttgt ttatgttggt ctttgtgtgt tgtctttgtt 360
gtgtttcttt tcttgtctct tttctgtttg tgtgatttat ctatttagtg tcgttgttgt 420
gctcctctgc cttcgtgctg cttctttcct cctgcgtcgt ctcgtctgtc ggtctacggt 480
gttgcgatat atccgcgatg ctgctgtcca atggcggacc taacgcgttg tgttaaccgc 540
ctagcgggtt ctcttcctct agttcaccgt atcggcttgg tataagagcg tgactggccc 600
tacggtgaac ctcgaaattg ttaaggggag cactctatac gtagagtcta ttgaacgttc 660
gcattgtaac tgctgtaacg ttatagctct ggataattcg cgcttcgatg atttgggaaa 720
gaaatgggaa tatcatcttg ttatgacgga agaaagggaa atcgagatcc tgaatatacg 780
ccatgatgcc gctgcttcta cgattgacgt ggcccaccta tttcggcaca actctaaagg 840
gggggcaccg ggttcaaacc gtgaaacggt gctatcgtaa aaatgggcgc ttccccaagg 900
gccttcgtgc aagtgcaata actctggcgg atcccgttat tgaccatgga aacctggcct 960
agaatcctcc taagtatcct caatgttact gtatgaacag gtagtcttcg tgttggtggg 1020
tatttatcgt agatcaagga cagtgagatg aatgcgacaa actgataagg ctgatcagac 1080
ctaggctcgt ttgatgcgac gaactacgtc tgcatacgag cgccgacaca tcttataccc 1140
gctatgagat ttaacaaggc ggcaggacga agccaaactc ctcagagcgc tcaccgcgct 1200
ctttccgcag cctccccccgc cgccggccgg gcgtagcgca gggtcacgtg ccggcgacgg 1260
ggcggggcca gcggggcgcc tgcgcgcctg cgcgccgctg gccgacggag gggagcctgc 1320
cgatgccgag cgggtgctac gtcccgcggt cggagccgcg tcttctcccg gctccgccac 1380
cagccggggc tcgggtgggg gcccggggcc ccggggcatg gccctccgga gtgcgcaggg 1440
cgacggcccc acctccggcc actgggacgg cggcgcggag aaggcagact ttaacgccaa 1500
aaggaaaaag aaagtggcag agatacacca ggctctgaac agtgatccca ctgatgtggc 1560
tgcccttaga cgcatggcta tcagtgaagg agggctcctg actgatgaga tcagacgaaa 1620
agtgtggccc aagctcctca atgtcaatgc caatgaccca cctcctatat cagggaagaa 1680
cctacggcag atgagcaagg actaccaaca agtgttgctg gacgtccggc ggtcattgcg 1740
gcggttccct cctggcatgc cagaggaaca gagagaaggg ctccaggaag aactgattga 1800
catcatcctc ctcatcttgg agcgcaaccc tcagctgcac tactaccagg gctaccatga 1860
cattgtggtc acatttctgc tggtggtagg cgagaggctg gcaacatccc tggtagaaaa 1920
attatctacc caccacctca gggattttat ggatccaaca atggacaaca ccaagcatat 1980

attaaactat ctgatgccca tcattgacca ggtgaatcca gagctccatg acttcatgca 2040
gagtgctgag gtagggacca tctttgccct cagctggctc atcacctggt ttgggcatgt 2100
cctgtctgac ttcaggcacg tcgtgcggtt atatgacttc ttcctggcct gccacccact 2160
gatgccgatt tactttgcag ccgtgattgt gttgtatcgc gagcaggaag tcctggactg 2220
tgactgtgac atggcctcgg tccaccacct gttgtcccag atccctcagg acttgcccta 2280
tgagacactg atcagcagag caggagacct tttgttcag tttcccccat ccgaacttgc 2340
tcgggaggcc gctgcccaac agcaagctga gaggacggca gcctctactt tcaaagactt 2400
tgagctggca tcagcccagc agaggcctga tatggtgctg cggcagcggt ttcggggact 2460
tctgcggcct gaagatcgaa caaaagatgt cctgaccaag ccaaggacca accgctttgt 2520
gaaattggca gtgatggggc tgacagtggc acttggagcg gctgcactgg ctgtggtgaa 2580
aagtgccctg gaatgggccc ctaagtttca gctgcagctg tttccctgaa agccagagaa 2640
gaccttcctc ttacatcaca ttaaggtctc acccttccat ggaaggattg ggaggggtgg 2700
aatttccttt attaaaaggg tttctgtcca gagtttttta ttcctgccac cctgccctgc 2760
ccgtttgtgt ttgcctttac ttcagaggcc aatagcagag cctacctgac acagagggga 2820
cctgggggcc tggcacagaa tcttatgcct ctcccagtgg gcttttggat gggctctctg 2880
ccacctccct ggtgttggag gaagtgggtt tctgaagctt gcagctgctt ttgcaacctg 2940
ctgggcctgg gctgccctag ttaccgggaa cctcctagtc accccataga aggtgaactg 3000
gaacatggga tgccctctta ctgctcctgc ttgcctgcat cagatggaca cagaaggagg 3060
gggctgctgc tgaaggagtc aggcttggga tggagagcaa gaggccaaag gactcgtctg 3120
ggagcatgat cggcccctct tactagcagc tccatcagct gatcagaaat cagagacaga 3180
agggaactag gttactcgtt aaatcctacc tattttcagc ctagaaatca catgggcgtt 3240
tccagtcaag tgggaaggaa gctcattttt agggcctaga cctagactag ttttgcctgc 3300
ctttttttcc tccctgcttt cttctacaag ccacacaaca ttgcttataa aaggtacttt 3360
tattactatt tttagaatat atacacccca tatgtaaatt accctcactt ttaaagtaat 3420
aatagttcct tcacggaaaa aaaaatagtt ctcatcagaa ggatgagctt ctaagatcct 3480
gaattttact ttctacttga ttaacatttt cgctttagcc agagttgcta tttgcaaaca 3540
gctttccact tccaaagtgc cactttctga ccggtttttc ttcctgtttt cctcaaatca 3600
tcctactccc aaataggcac ccactcacta ccttggcgtc gtttttttggg tctgacttgt 3660
ctcgtcaact gctggtctct cccacctggt tggaaatgtc gttggaaact tgcaaagact 3720
ctccagacct tagggaacaa gaggcatcac tcagtccttc tgggacagct tccctgtaag 3780
caaagccagg tttgttttag attcactctg gcctgaaaaa aaaagtgcct tttgctgctt 3840
taaagaattg ggtatatgg tatgaagcag ccatgtactt gtattttcct ggtctttcct 3900
gggcactctt ctctcttggc agatgttttc ttaaagtgaa cacaccagaa gcactctacc 3960
ccaccatatc cagtcctctt tgcatgcctg ttgtgtgcag ggtggagggt gtcccacgtg 4020
ccagagcagc tgggagccct gtccgctgag ttacactaag cactttattt ttatttattt 4080
tattattatt ttcttttttc tgagatggag catgggcaaa aagagcgaaa gcta 4134

<210> 70
<211> 2329
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 677977CB1

<400> 70
atgaaacatt gccttccgaa aagcctacta catccaagga caagttgcag gccatcgagc 60
cgcgtcaatg ggcggcgagc gcaaggccca gacgtgtgcc gcgacattca gcgtgccggc 120
gcgggcatgc gcggctggct caaggacgat gcccacatgc gctggttcct ggagcagctg 180
ggcggtgagg tgggcactca gcgcaagaag atgcggctgg ccaacgagga ggagatcgac 240
cgcgccgtgt acgcctggtt cctggcgctg cgccacgacg ggtgccgct gtctggcccg 300
ctcatccagg cgcagccgag gccttcgcgc gccagatcta cgggcccgag tgcaccttca 360
aggccagcca cggctggttc tggcgctggc agaagcgcca cggcatctcc agccagcgct 420
tctacggcga ggcgggcccc cagccccgag ccccgcgccc ggcccgcccg tcaaggagga 480
gcccgcgctg ccctccggcg ccggccccct gcccgaccgc gccccggccc cgccgccccc 540
ggccgagggc ggctacgggg acgagcagat ttacagcgcc agcgtcaccg gcctctactg 600
gaagctgctt ccggagcagg ctgcgccccc gggcgcaggg gaccccgggg cggggggctg 660
tggccggcgc tggcggggcg accgcgtaac ggtgctgctg gccgcaaacc tgaccggcag 720
ccacaagctg aagccgctgg tcatcgggcg gctgccggac ccgcccagcc tgcgccacca 780
caaccaggac aagttcccgg cctcctaccg ctacagcccc gacgcctggc tcagccgccc 840
gctgctgcgg ggctggttct ttgaggaatt tgtcccaggc gtcaaacgct acctgcgccg 900
aagctgcctg cagcagaagg ccgtgctgct ggtggcccac ccgccctgcc caagcccagc 960

tgccagtatg cccgccctgg acagcgagga tgccccgtg cggtgcaggc cggagcccct 1020
cggtcccccg gaggagctgc agacaccgga tggcgctgtg cgggtgctgt tcctgtccaa 1080
aggcagcagc cgggcacata tccccgaacc ggtggagcag ggcgtggtgg ccgccttcaa 1140
acagctgtac aagcgcgagc tgctgcgact ggctgtgtcc tgcgccagcg gctccccgct 1200
ggacttcatg cgcagcttca tgctcaagga catgctctac ctggctggcc tctcctggga 1260
cctggtgcag gcgggcagca ttgagcgctg ctggctgctg ggcctgcggg ctgccttcga 1320
gccccggccc ggcgaggaca gtgctgggca gccggcccag gccgaggaag ccgccgagca 1380
cagcagggtg ctcagcgacc tcacccacct ggcggctctg gcctacaagt gcctggctcc 1440
ggaggaggtt gcggagtggc tgcacctgga cgatgatggg gccagtctgc cctctgccat 1500
gggggcgga gaggacgagg aggaggccac cgactatgga gggacctcca gtctgccctc 1560
tgccattggg ggcggagagg acgaggagga ggccaccgac tatggaggga cctcagtgcc 1620
gactgccggg gaggccgtgc gggggctaga aacagctctg cggtggctgg agaaccagga 1680
ccccagagag gtgggggccac tgaggctggt gcagttgcgc tcactcatca gcatggcccg 1740
gaggctgggg ggcatcgggc ataccccagc aggcccctat gacggtgtgt gaccaggcca 1800
gcccagtgac ctttctcctg ctgcacttgg agggagggga catacacaca gtctcccatc 1860
tctcctcccc tccccctggg gtggcccacc gcatgggtac agggggttcc aggaatccaa 1920
atccagcatg gcttggagga gctctgttgg tgagaggtcg ccctgcctca ctggcaccct 1980
gggggcacag ctggaagaga ggcctggccc atgctcctct cagggcaggc acatgtacgg 2040
ggcatacaag gcacagcgcc tgttggaaca ggtggctgtg ttcctgctct ggcccccgtg 2100
cggctgggcc tccgcccctt gcaccagtca catgcactgg acgagggccg aaactcctgt 2160
ctgctatcga gcctggtgct atgtggcccc ggagccacag caaatcatct cgagtggcga 2220
ggcacccact tgatctcttt ttcttaaccc tttatctggt tcaggttaca caaaaagggg 2280
gcgcccgatt gggccgcccg gaaaatccgc gcccggggac gtccttggg 2329

<210> 71
<211> 2960
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1661472CB1

<400> 71
tcaagttctg attctgagct gctccgccac tggcctgtgg gaagaggcag gaggcttggg 60
gtcccccagc ctgtgcccag cctcctgtgc cttgtagatc ttcacatatg agcgcaagtg 120
gtactcgcgc aaggacctgg cccggcatcg catgcagggt gaccccgatg acacgtcgca 180
ccgtgggcac ccgctctgca agttctgtga cgagcgctac ctggacaatg atgagctgct 240
taagcacctg cgccgcgacc actacttctg ccacttctgc gactcggacg gggcccagga 300
ctactacagc gactatgcct acctgcgtga gcacttccgg gagaagcact ttctgtgtga 360
ggaaggccgc tgcagcacgg agcagttcac ccacgccttc cgcaccgaga tcgacctcaa 420
ggcccacagg acggcctgcc acagtcgcag ccgcgccgag gcacgccaga accgccacat 480
cgacctgcag ttcagctacg cgccacggca ctcgcgccgg aacgaggggg tcgttggtgg 540
cgaagactac gaggaggtgg acaggtacag ccgccagggc cgagtggccc gggctggcac 600
tcgcggagcc cagcagagcc gccgaggaag ctggaggtac aaaagggaag aagaggaccg 660
agaagtagca gctgctgtcc gggcctccgt ggccgcacag cagcaggagg aggctcgcag 720
gagtgaggat caggaggaag gcggtaggcc caagaaggag gaggcagcgg cgcggggacc 780
tgaggatccc cgtggccccc ggcgctcacc ccggactcag ggcgaaggcc caggccccaa 840
ggaaacctcg acaaatggtc ctgtaagcca agaagccttc tcggtgacag gcccagccgc 900
cccagggtgt gtgggggtgc caggcgccct cccaccaccc agcccgaagc tcaaggacga 960
agacttcccc agcctctctg cctccacttc ctcctcctgc tccactgcag caaccccggg 1020
ccctgtgggg ttggcgctgc cgtacgccat ccctgccaga ggcaggagtg ccttccagga 1080
ggaggacttc cccgccctgg tgtcctcggt gcccaagcct ggcaccgccc ccaccagcct 1140
tgtctctgcc tggaacagca gcagtagcag caagaaggta gcacagcccc cactctcggc 1200
gcaggctacc ggcagcggcc agcccaccag gaaggctggg aaggggagca ggggcggcag 1260
gaagggcggc ccgcccttca cacaggagga ggaggaggac ggcggcccgg ccctgcagga 1320
gcttctgagc acacgcccca cggctccgt ctcctccaca ctggggctgg cctccatcca 1380
accctctaaa gttgggaaga agaagaaagt gggctcggag aagccaggca ccacattgcc 1440
acagcccccg cccgctacct gtcccccagg ggctttgcag gccccggaag ctcctgccag 1500
cagagccgag gggccagttg ccgtcgtcgt taatggacac acggagggcc cggcccctgc 1560
tcggagtgcc cccaaggaac cccctgggct cccaaggccc ctggggtcct tcccctgccc 1620
cacgccacag gaggacttcc cagcgctcgg cggcccctgc ccaccccgga tgccgccgcc 1680
cccaggcttc agcgctgtgg tgctcctgaa gggcacgcct cccccacccc cgccgggcct 1740

ggtgccccca atcagcaagc cgccccctgg cttctctggc cttctgccta gcccccaccc 1800
ggccagtgtc cccagccccg ccaccaccac caccacaaaa gcacccaggc tgctgcctgc 1860
cccacgggcc tacctagtcc ccgagaactt ccgggagagg aaccttcagc tcatccagtc 1920
catcagggac ttcctgcaga gcgacgaggc ccgcttcagc gagttcaaga gccactcagg 1980
ggagttcaga cagggcctga tctccgcagc ccagtattac aagagttgcc gggacctgct 2040
gggggagaat ttccagaagg tctttaatga gctgctggtc ctgctgcccg acacggccaa 2100
gcagcaggag ctcctgtctg cacacacgga cttctgcaac cgcgagaagc ctctgagcac 2160
caagtccaag aagaacaaga agagcgcgtg gcaggccacc acccagcagg cgggcctgga 2220
ctgccgtgtg tgccccacct gccagcaggt gctcgcgcat ggcgatgcca gcagccacca 2280
ggcgctgcat gctgcccggg acgacgactt cccctccctg caagccatcg ccaggatcat 2340
cacgtagctc ccgccagcgt ggccagagct gtcgcaccgt gagcgtcctt cctccttcct 2400
ctccgggctg ccaggcagcc aggtaaggcc tggtgaggcc acttggcctc ttggttggcc 2460
aggcccacca ggaagtcacc aggacagtcc acccgccctg ttggcacact caagcgggag 2520
tccacccctg cctcagtggt gggccagtct cggtttgcat tcttgtgctt ttgggaggtg 2580
ccaggggagg gaagggctgg gatgctggga cctgttgttg ctggcaaagc cagaggtcac 2640
agtggcctga tctgggccct cccaaagctg agggctgcag cccgtggggc ctcagagctg 2700
aaagctgcgg cgccactggt gccagagtca gatgtcacag atgtgttgtg taaacagttg 2760
gctgtttcgt gcttcaagaa tgttcaggat taaaagcaga caagaaattg tgctacttga 2820
agttgaatct ttttatgaga caagctgaat ctgggatctc aaattgcctc tgacctttta 2880
taagacagtt tatcttcaaa taaatttatt ttgcaatacc acgcaaaaaa aaaaaaaaaa 2940
aaaaaaaaaa aaaaaaatat 2960

<210> 72
<211> 2623
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1748508CB1

<400> 72
tcagttgatg tgttaggaaa attgattata tgtatatact aaataaaatt tgttaaacat 60
acctataacc cttcaaaatt aaatttttaa aaacttttta ctgtttcatg taggtataac 120
ttacttcctt gtctcttgat agatttcagg cttaaactga agaagtaatt ggttcaaaaa 180
tcttattact agaaagtaat gacaaattta tagttttcat ttctagtgta cattaaccat 240
aggtcattgt aacatccctt ctccacaata aaaatggaaa cttattttgt ggaaatcatt 300
ttatgtaaat atgtatttaa tacttatttc atctttctta cctttcagaa ttaccatgaa 360
attatgactc gtcatcctga gaattatcaa tgggaaaact ggagtctaga aaatgttgct 420
accattttag cccaccggtt ccccaatagt tatatttggg tgataaaatg ttcccgaatg 480
catttgcaca aattcagctg ctatgacaat tttgtgaaaa gtaacatgtt tggtgcccca 540
gaacacaata ctgactttgg agcttttaag cacctttata tgttattagt taatgctttt 600
aatttaagtc agaatagttt atcaaagaaa agtttgaatg tttggaataa ggactccata 660
gcatctaact gtagatccag tccttctcat actacgaatg gttgccaggg agaaaaagtg 720
aggacctgtg aaaaatctga tgagtctgcc atgagttttt atccaccatc actaaatgac 780
gcatctttta ctttgattgg attcagtaaa ggttgtgttg ttttgaatca gttgcttttt 840
gaattgaaag aagccaagaa agacaagaac atagatgctt ttatcaaaag cataagaaca 900
atgtattggc tggatggtgg tcattctgga ggaagcaata cttgggttac ttatccagaa 960
gtcttgaaag aatttgcaca aacaggaatt atcgttcaca ctcatgtaac accttaccaa 1020
gtacgtgatc caatgagatc ttggattgga aaggagcaca agaaatttgt tcagatactt 1080
ggggatcttg gtatgcaggt gactagccaa attcatttta caaaggaagc tccttccata 1140
gagaatcact tcagggttca tgaagtattt tgagattaca ggtatattaa tgaacttgtt 1200
cagtggaaga acataagcac ttttgagtgt tataaattca gataatggga tgtaattcat 1260
agctgcattg tcagttttgg ggtatggggg gaagcacaca ttcctaaaat gtgagtgtaa 1320
tgtgcaatag tatttttttgc ttgtgaatgt gagcagttat taatttggat tgagttagaa 1380
ttagttaatt tgaaatctaa caaggtggtt tgtaataatg ctgaggagat ataagaccct 1440
taaaatgaaa gttacaacat tgttcttata aaaggtaact aaaattgtta ctgttggaaa 1500
taactgattt tctgagtaat gttttaaact aatttggtga cattttaaca gtaattagct 1560
attttgagtg gaaatatttt catttctctt caaacaaaag caaaggtacg atgctgtttt 1620
ctatcatttt ggaataactg caccctgcct tttgtgtttt tgtaaactcc ttgactcatt 1680
ctttcatgtg tcaccaagta cttttctcat gagagtcaac atatatttgt ttccaaatgt 1740
ccacaagtgt acaatagtgt aaaggtggtt tttaaaaaca tagccaggtg tggtggcacg 1800
tgcctttagt tccagctact caggaggcta aggcaggagg attgcttgag cccaggctgt 1860

gtggttcacc ataattgtgt ttgtgactag ctactgcact ccaacctggg caacatagtg 1920
ggacttcatc tctaaaacaa aacaaaacaa aattacactt aagcactatt gtttaatttt 1980
taattgtcag tttatcatta ttttgggtaa gacattctgg ggtttcttga atcttgtcca 2040
aaaaccagtt gttttggaaa attgctttaa attgagcata tttatgtata ttggataaaa 2100
atgtactaca gagcaaattt caaatttttc attatatcag tctttttgaa aggatcaact 2160
tggataaaat aaatatataa tgctctattt gttagagctc tattaaaaag gaaacagatt 2220
ccatagatct aagtcaatgt ttctccagaa gcatgatttt gtctgccaaa agaaaatagc 2280
tctctttggc caaaatgcaa aattacattg ctataagaaa agttacaagg gaaagtttga 2340
agacacaaat gatttaattt tggctcaaaa actgaatttg cttaacactg ctacataatt 2400
tgggtgaagt ttccttctgc ccgtttttct tgacctagat aaatacactt tgagaaatcc 2460
agatctaata aatgtcaacc aacattgaca ttgtaattgg gtgattacaa taaaaggtga 2520
gcagtttgtt gtttattaat aattagcttt tgcaggtaat gaaatagcag ggaagtaaca 2580
tgctgcttta ggactaaaaa aaaaaaaaaa aaaaaactcg gtc 2623

<210> 73
<211> 4518
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2159545CB1

<400> 73
gtggctgcac cttagcgaca tcaggaagaa cagggtcgag gatcgagtca cagcaaactg 60
tacaatggca gttgagtggg ttttgaaaaa cgtctgaatt gtagtgtcta catctacctt 120
ctgccttctc tctacccttt gaaaaaaaat aaatggcatg tttgagaaga gagctctatt 180
aaaagaggaa tattgcaaat ggcttcagat tctatgagta gtaaacaagc taggaatcac 240
attacaaagg ggaaaaggca acaacagcac cagcaaataa agaacagatc ctcaattagt 300
gatggtgatg gagaagattc ctttattttt gaagcaaatg aagcttggaa agattttcat 360
ggttctcttc ttcgatttta tgaaaatgga gaactctgtg atgtcacact caaggttggc 420
tcaaagctaa tctcttgtca caagctggta ttggcttgtg ttattcccta ctttagagcc 480
atgtttcttt ctgaaatggc tgaagccaag caaacgctga ttgagattag agattttgat 540
ggtgatgcaa tagaagactt ggtaaagttt gtctattctt cacggctcac tttgactgtt 600
gacaatgtcc agcctctctt atatgcagcc tgtattctgc aggttgaact ggtggctaga 660
gcttgttgtg aatacatgaa gttacatttt catccctcca attgcctggc agtaagagcc 720
tttgcagaaa gtcacaatcg aatagactta atggacatgg cggatcagta tgcctgtgac 780
cattttactg aagtagtgga gtgtgaagac tttgtaagtg tatcaccgca gcacctccat 840
aagcttttgt cctccagtga tctaaatatt gaaaatgaaa agcaggtcta taatgctgcc 900
atcaagtggc ttcttgccaa tcctcagcat cattccaaat ggttggatga aacacttgca 960
caggttcgcc tgccattgtt gccggttgat tttcttatgg gtgttgtggc aaaagaacag 1020
attgtcaagc aaaatctaaa atgtagagat ttactggatg aagcaagaaa ttaccacctt 1080
cacttgagta gcagagcagt acctgacttt gaatactcca ttcggactac cccaaggaag 1140
catactgctg gtgtgctgtt ttgtgtaggt ggtcgaggtg gatctggtga cccctttcgc 1200
agtattgaat gctattctat caacaaaaac agttggttct ttggaccaga aatgaatagt 1260
cgaaggcgac atgtgggtgt aatctctgtg gaaggtaaag tgtatgcagt aggtggacat 1320
gatggaaatg aacatttagg gagtatggag atgtttgatc ctctcactaa taaatggatg 1380
atgaaggcat caatgaacac aaagaggcga ggaattgcct tggcttcctt aggaggccca 1440
atttatgcaa ttggagggtt agatgacaat acttgcttca atgatgtgga gagatatgac 1500
atagaatctg atcagtggag tacagtggca ccaatgaata ctccccgtgg aggagttggc 1560
tctgttgctc tagtaaacca tgtttatgca gtaggtggca atgatggaat ggcttcttta 1620
tctagcgtgg agagatatga tccacatctg gataagtgga tagaagttaa agaaatgggt 1680
cagcgaagag caggcaatgg agttagcaag cttcatggtt gcttatacgt agttggtggt 1740
tttgatgata attctcctct gagttcagtt gagcggtatg acccccgaag caacaagtgg 1800
gattatgtgg cagcacttac tactcccaga ggtggagtgg gaatcgcaac agtgatgggc 1860
aaaatctttg cagttggtgg tcataatggc aatgcatact taaatacagt agaagcgttt 1920
gatccagtgc tgaataggtg ggagcttgtt ggatctgtgt ctcactgcag agctggagca 1980
ggagtagctg tgtgttcctg tttaactagc caaattcgag atgtaggtca tggatccaat 2040
aatgtggttg actgtatgtg atttgagttc tggccaccaa caatttagtc atatctgata 2100
ggtacaaaag aaaaccaaga ttttgatatg accacctttc aacacttttt actgcaacta 2160
gagtcttatt taaatgttaa ctgttgtatt gtgactaagt acaacttttg gaatgatacc 2220
tgaagaatta ttagtaatag agttaaattt gatacttcct gcttttgcaa ataatggagt 2280
agggaagaag atagtcgaac ttaaaatcat acctaccctg aaaaatgctt aaagtagtgc 2340

caaaaatctt gaaatctctc tctctctctc tatttggatg aatcaaaatg taatttgttg 2400
agttttgcct tttcttttgc ctaaagatgt taacgtcagg cagtgggttt ttcctagagg 2460
agagattagt acctattcat taacacaagg atacataaga gacttgcccc ttttgaagat 2520
catttctact tggtgctgat gaatttatgc ctagcttttt ttttctaggt ggtgattttt 2580
caatttcttt tatatttacg gaagtagaaa aaatgttctg tctttttgaa tgccacttag 2640
ctgcattttc agaagcttat aggctataag aagcctttaa tttcagtttg atttgctagg 2700
acaccaaaag acctagctgg gctaactcca tccaaatgat tttacagttt tattttttca 2760
tcttttgcaa aaagtatttg ttacataatc aaaaggtgaa ccagaggaaa tattgaacaa 2820
ggttttagat ctgatgttat tctggtgccc aattgagaga ccttgagaaa gaatgatgca 2880
gattacatga ctctaatgtg ggcagtctga aaatatcttt aactgtcttc tttgatcatt 2940
tgattttgga tccaagattt atcatctctc tattttgcta tttgccacca tcctggtttt 3000
tctgactgta cagaatcaat ctgcatttat ttattatcag aaatgaatgc atttttacaa 3060
aaaccttttt gtagactgta gtttattgat tagttcaagc ccatcacgaa ctaattaaac 3120
atattgttgg taaaaacaga gattgaaatt actgttttta tcatatttaa atccctctgg 3180
tgtatttatt gaaatttaat taaaaatatt tctacatatt gatttgaagc caggatattg 3240
atggacctac agatttaaca gctcttcaaa gttttaacta gcattataaa atggcctatt 3300
cattatatgt atattttttt ataaatattt atcaagaaca tgaaaatatt ttattttgta 3360
tttgcagtgc agtttggcac ctctcttcca ggccttccta agtttttgtt agttgtgtct 3420
tcgaatttat tgtagttcac cgctggtcag taatctggtt cataaggaag ctggggcagt 3480
gaaaaacact tcgttgatcc atccagccct aaagttcatt gactaatagc tcatgcatct 3540
tattaaatgg gatttggac ttctacaata gtgtggaaag ccacatcaca aaaaaagcca 3600
catcacaaaa aaaagccagt gcctgttgag tatatcaaac caaaaaagaa aaaaaaaagt 3660
tttaagaggg gtcaggctgg cagataaaag tctcttatga agcagatata aatccagtaa 3720
gctatttttt tctacctatg ggcaagttca taagtcattg ctggatctga accagtgttc 3780
ttttgtgtaa gtacaaccat atttcagtca gtccagaggc agttaaccct gtagcactta 3840
ctagactggt gacaactggc tggatgtgga gaataagtgt caaaacctgg cttcaatgtc 3900
tatttcttat gtctaatgga gttcttatat tacagtgttt gtttagctat ggttactaat 3960
atgtttgtta tatttatttt cggggagagg atcttgttat tgtttcctga gattatttaa 4020
aatgtttttt ctttgatgta tttctgtctt tcttatatgt aaggagttat aagggttatt 4080
atattttcat ttctggtttt tactatacaa tggttatcta acatactaag ttggtaccct 4140
acctacccgt aaaatatgaa atactcagaa gtaaagtgtg tgggcaatca gaagcatagc 4200
cttgaaggac ctgacagcag tatgaaaaga taaataggcc aggcatggtg gctcacgcct 4260
gtgatcccag cactttgggg ggccgaggtg ggtggaccac ctgagatcag gtgttcaaga 4320
ccagcctgac caacatggtg aaaccctgtc tctactaaaa gtacaaaaat tagccaggcg 4380
tggtggcgca tacctgtaat tccagctact caggaggctg aggcaggaga attgcttgaa 4440
cccaggaggt ggaggttgca atgagccgag gttgagccat cgcactccag cggtgaaaca 4500
agagtgaaac tccatctc 4518

<210> 74
<211> 1238
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8560269CB1

<400> 74
ccgcccccgc gtcgcgggct ctttaaggcc ggcggttttc ggcagccagc ccggggcggg 60
gaaagcggag cgcgcgctcc acgcgggacc gcctcccggg ccgtctgagc agagggcggg 120
gtgcaggcgg aatggccctc gtgccctatg aggagaccac ggaatttggg ttgcagaaat 180
tccacaagcc tcttgcaact ttttcctttg caaaccacac gatccagatc cggcaggact 240
ggagacacct gggagtcgca gcggtggttt gggatgcggc catcgttctt tccacatacc 300
tggagatggg agctgtggag ctcaggggcc gctctgccgt ggagctgggt gctggcacgg 360
ggctggtggg catagtggct gccctgctgg gtgctcatgt gactatcacg gatcgaaaag 420
tagcattaga atttcttaaa tcaaacgttc aagccaactt acctcctcat atccaaacta 480
aaactgttgt taaggagctg acttggggac aaaatttggg gagtttttct cctggagaat 540
ttgacctgat acttggtgct gatatcatat atttagaaga aacattcaca gatcttcttc 600
aaacactgga acatctctgt agcaatcact ctgtgattct tttagcatgc cgaattcgct 660
atgaacggga taacaacttc ttagcaatgc tggagaggca atttattgtg agaaaggttc 720
actacgatcc tgaaaaagat gtacatattt acgaagcaca gaagagaaac cagaaggagg 780
acttataatt ggctataatt tataagaatg ttgtcattga gtgtgtcact taaggtctta 840
gactgcaaat ctaaccatat ttaatgaaat gtcttactgt acaaaaagtc taagccaaag 900

gttctcaggg gagaaagcac atgtgcagtt ttaaaacaaa gcagtgcttt gtcccattgc 960
tgtgattttt agtcagactt tactcagtct gaaatgcaat taacattaaa ggattaagtg 1020
tgagatttcg atttatgcta tttgtgtatc ccatactcct cccttttaat aaacagtttc 1080
cactgatgat atgaagggcc ggtataaaga agtctttaaa tgagtaagct ttcttggtaa 1140
gattaaatct tacaaattat ttttaaaacc ttgtgatata tacaatgtta gctgagtttt 1200
ctaattttct ggatgtaaaa caaaaggttt aacctata 1238

<210> 75
<211> 1771
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8710302CB1

<400> 75
gcgatacgcc gagcgctcaa catccgagga ctttggccca gagtaacccc gctctcgtga 60
cctttcccct ccattccgca cctccgagtg ctggccgggc gagaggctgg cggctgggct 120
ctcgcgcccc tccctgcagg gctcaggctc tccccctcct gtcttctccg cgctgttcct 180
cgtcatggcg gccctcagca agtccatccc tcataactgc tatgagatcg gccacacttg 240
gcaccccttcc tgccgggtct ccttcctgca gatcaccggg ggcgccctgg aggagtccct 300
gaagatctat gctcctctgt acttgattgc agcaattctc cggaaacgga aattagacta 360
ttatttacac aaactactcc ctgagatcct acaatccgct tcatttctaa ctgctaatgg 420
ggccttgtat atggctttct tttgcatttt aaggagaggg ctgctcacaa tttatatggc 480
caacttggcc acagaaacac tattcagaat gggtgtagca agaggaacca tcacaacatt 540
aagaaatgga gaagtccttt tgttttgcat cacagctgcc atgtacatgt tctttttcag 600
gtgcaaggat ggcttgaaag gatttacatt ttctgcactt aggttcattg tagggaagga 660
agaaattccc acacattctt tttcaccaga ggcagcatat gcaaaagtgg aacaaaagag 720
agagcaacat gaggaaaaac ccagaagaat gaatatgatt ggtctagtca ggaaatttgt 780
ggattcaata tgcaaacatg gaccaaggca tagatgttgc aaacattatg aagataattg 840
catctcttat tgcattaaag gtttcatcag aatgtttagc gtggggtact tgatccagtg 900
ctgcctccga atcccttctg catttaggca tctgtttaca cagccatctc ggctactttc 960
tctcttctac aataaagaaa acttccagct tggagctttt cttggctctt ttgttagtat 1020
atacaagggt actagttgct tcctgcgctg gatcagaaac ttagatgatg aactacatgc 1080
tattatagct ggattttttgg caggtatatc aatgatgttt tataaaagca caacaatttc 1140
catgtattta gcgtccaaat tggtagagac aatgtatttc aaaggcattg aagcagggaa 1200
ggttccctat tttcctcatg cagatactat catctattcc atctctacag caatttgctt 1260
ccaggcagct gtcatggaag ttcagacttt gagaccatct tactggaagt tccttttaag 1320
actcaccaag ggcaaatttg ctgtcatgaa ccgaaaagtc cttgatgttt ttggtactgg 1380
tgcatctaaa cactttcagg atttcatccc caggttggat ccaagataca caactgtaac 1440
accagagttg cccacagagt tttcctgaag atgactgtaa cttattaatg tgactaaatg 1500
tttcatcttg aagagttaat tatgttgaac acaaaggagg gggcccaagc tcgaacttca 1560
gtgttatttc agttagagat actcttttca tttgttttgt ttttcttatg aatcagaaat 1620
tcagaagctt tttaggaaga tgttgcttaa taattaagct tcctccatag ccagaataag 1680
attctggatc actgtagtga ctgacattat atattattga tcaaattatg tccacaagca 1740
atattatata atctacgtag aagtgtaata a 1771

<210> 76
<211> 2909
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6778214CB1

<400> 76
tttccgtggc ccttcctgcc ctttgtttct ctcttgatgt cacccgcggg acaactctac 60
gtgcaagtcg ctgaaatcga ttttctgctt cttatagtaa gcggcgggct cgccagcctc 120
gagcgaaagt gtgactgcga acgggcaggc gcgcgcgggg ctcggcggag gcgcgcttgg 180
gctcccggcg gcgacgacta cgaccactag gagagcggac ggaggcggcg cctgaagcgg 240
cggcggagcc catgccccgg gacggcgggc ggacccggag agacaaatcc ggggcccggg 300

gcatgtcccc ggggcccccg tgaggaggcg gcggcggcta tggagatcgc gccgcaggag 360
gcgccgcccg tgccgggcgc ggacggcgac attgaagagg ccccagctga ggccgggtct 420
cccagccccg cgtcgccccc cgccgatggg cgcctcaagg ctgcagccaa gcgcgtcaca 480
ttcccgtccg acgaggatat cgtgtctgga gcagtggagc ccaaagaccc ctggagacat 540
gcccagaatg tgaccgtgga cgaggtcatc ggcgcctaca agcaggcctg ccagaagctg 600
aactgcaggc agatccccaa gctcctcagg cagctgcagg aattcacaga cctcgggcac 660
cgcctcgact gtctggacct gaaaggtgag aagcttgact acaagacctg tgaggccctg 720
gaagaggtct tcaagaggct gcagttcaag gtcgtggacc tggagcagac aaacctggat 780
gaagatggtg cctcggccct cttcgacatg atcgagtact acgagtcggc cacccacctc 840
aacatctcct tcaacaagca catcggcacc cggggctggc aggcggccgc ccacatgatg 900
cgcaagacga gctgcctgca gtatctggac gcccgcaaca cgcccctgct ggaccactcg 960
gcgcccttcg tggcccgtgc cctgcgcatc cgcagcagcc tggcagtgct gcacttggag 1020
aacgccagcc tgtcggggcg gcccctcatg atgctcgcca cggccctgaa gatgaacatg 1080
aacctgcggg agctgtacct ggcggacaac aagctcaacg gcctgcagga ctcggcccag 1140
ctgggtaacc tgctcaagtt caactgctcc ctgcagatcc tggacctccg gaacaaccac 1200
gtgctagact cgggtctggc ctacatctgc gagggcctca aggagcagag gaagggctg 1260
gtgaccctgg tgctgtggaa caaccagctc acgcacacag gcatggcctt cctgggcatg 1320
acactgccgc acactcagag cctggagacg ctgaacctgg gccacaaccc catcgggaac 1380
gagggtgtgc ggcacctcaa gaacggctcc atcagcaacc gcagcgtgct gcgcctcggg 1440
ctggcctcca ccaagctcac gtgcgagggc gcggtggcgg tggcggagtt catcgctgag 1500
agcccccgcc tcctgagact ggaccttcgg gagaacgaga tcaagacagg cgggctcatg 1560
gcactgtcgt tggccctcaa ggtgaaccac tcactgctgc gcctggacct cgaccgtgaa 1620
cccaagaaag aggcggtgaa gagcttcatc gagacgcaga aggcgctgct ggccgagatc 1680
cagaacggct gcaagcgcaa cttggtgctg gcgcgggaga gggaggagaa ggagcagccg 1740
ccacagctgt cggcctccat gcctgagacc accgccaccg agccccagcc cgacgacgag 1800
cccgccgctg ggtgcagaa cggggccccc agccccgcac ccagcccgga ctcagactca 1860
gactcggact cggatgggga ggaagaggag gaagaggaag gggagaggga cgagaccccc 1920
tccggggcca ttgacacccg ggacacaggg tcctctgagc ctcagccacc accggagccg 1980
cctcggtcag ggccaccact gcccaacggc ctgaagcccg agttcgccct ggcactgccc 2040
cctgagccgc ccccggggcc tgaggtcaag gggggcagct gcggcctgga gcacgaactg 2100
agctgctcca agaacgagaa ggagctcgag gagctgcttc tggaagccag tcaggaatcc 2160
gggcaggaga cactgtgaca ctttagttcc ctttttccgg tcggtctgcg atgagctgag 2220
gccagagcca tgagaatctg ctcaccttcc ccccagcctt cctgaggccc aggatgccag 2280
gggtgggggc cattctgggg ccccccctccc cccacagcaa cactacaagg ggtgcaggag 2340
ctacagggag tggccctccg cgcgtgactc aagcacttct atttatgagc ccagcactgg 2400
aagactctgg gggtgaatgg gaggaggggg agcaggagga ggaggaggtc tccaagtaca 2460
tcaggcgcct gttctggagg ggccaggctt gccctgcgga gggcaggcgt cctgggtggt 2520
ggtgggatgg tcccctgtgg ccccgggcac agggccgggc aggcagcctg gtgccggaga 2580
ggcggtgcgt gctggtggtg gttgagatgc gcagaacagc cccagacagc gcaggccggg 2640
cagggtgggg ggatgggagc agaggatcag agctttcttt ttctcaagtg caataaatct 2700
atcagggagc tggggcggga gcagccggca ttccgggacc ctgttgtcca ggccactgga 2760
ggctgcgccc tgagaggcac tacagccctt tgggggcgag tggcatgggt gtgggtgagg 2820
gtgggcagag ggctggggcg actcctgtcg gtgcaactct gttcacacct tttagaataa 2880
actggcatgg tcggtaaaag acaaatagc 2909

<210> 77
<211> 2216
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 258383CB1

<400> 77
gcgatacgcc gagcgctcaa catccgagga ctttggccca gagtaacccc gctctcgtga 60
cctttcccct ccattccgca cctccgagtg ctggccgggc gagaggctgg cggctgggct 120
ctcccgcccc tccctgcagg gctcaggctc tccccctcct gtcttctccg cgctgttcct 180
cgtcatggcg gccctcagca agtccatccc tcataactgc tatgagatcg gccacacttg 240
gcacccttcc tgccgggtct ccttcctgca gatcaccggg ggcgccctgg aggagtccct 300
gaagatctat gctcctctgt acttgattgc agcaattctc cggaaacgga aattagacta 360
ttatttacac aaactactcc ctgagatcct acaatccgct tcatttctaa ctgctaatgg 420
ggccttgtat atggctttct tttgcatttt aaggaagata cttggaaaat tctactcatg 480

gactcctggc tttggtgccg ctctgccagc atcttatgtg gccattctca ttgaaagaaa 540
aagcaggaga gggctgctca caatttatat ggccaacttg gccacagaaa cactattcag 600
aatgggtgta gcaagaggaa ccatcacaac attaagaaat ggagaagtcc ttttgttttg 660
catcacagct gccatgtaca tgttcttttt caggtgcaag gatggcttga aaggatttac 720
attttctgca cttaggttca ttgtagggaa ggaagaaatt cccacacatt ctttttcacc 780
agaggcagca tatgcaaaag tggaacaaaa gagagagcaa catgaggaaa aacccagaag 840
aatgaatatg attggtctag tcaggaaatt tgtggattca atatgcaaac atggaccaag 900
gcatagatgt tgcaaacatt atgaagataa ttgcatctct tattgcatta aaggtttcat 960
cagaatgttt agcgtggggt acttgatcca gtgctgcctc cgaatccctt ctgcatttag 1020
gcatctgttt acacagccat ctcggctact ttctctcttc tacaataaag aaaacttcca 1080
gcttggagct ttcttggct ctttgttag tatatacaag ggtactagtt gcttcctgcg 1140
ctggatcaga aacttagatg atgaactaca tgctattata gctggatttt tggcaggtat 1200
atcaatgatg ttttataaaa gcacaacaat ttccatgtat ttagcgtcca aattggtaga 1260
gacaatgtat ttcaaaggca ttgaagcagg gaaggttccc tatttcctc atgcagatac 1320
tatcatctat tccatctcta cagcaatttg cttccaggca gctgtcatgg aagttcagac 1380
tttgagacca tcttactgga agttcctttt aagactcacc aagggcaaat ttgctgtcat 1440
gaaccgaaaa gtccttgatg tttttggtac tggtgcatct aaacactttc aggatttcat 1500
ccccaggttg gatccaagat acacaactgt aacaccagag ttgcccacag agttttcctg 1560
aagatgactg taacttatta atgtgactaa atgtttcatc ttgaagagtt aattatgttg 1620
aacacaaagg agggggccca agctcgaact tcagtgttat ttcagttaga gatactcttt 1680
tcatttgttt tgtttttctt atgaatcaga aattcagaag cttttagga agatgttgct 1740
taataattaa gcttcctcca tagccagaat aagattctgg atcactgtag tgactgacat 1800
tatatattat tgatcaaatt atgtccacaa gcaatattat ataatctacg tagaagtgta 1860
ataacaaaca agagtacact taaaattact ttaaaagatg tctttagttc attccaatat 1920
aattcttgat taaaattagg attatttcta cattttagga tttacaaagg atcacgggta 1980
catggatttg gtctatatat ttttttaaag ttttgaattg gtatctgtag tagtggaatg 2040
ttatagattt gaagtaactc tccacggaca gtgctgcttt cgtgtagagc aatttaattg 2100
gagaagtggc cattcttact tcagggatgc aaagatgggt ctcataccat ttggataaat 2160
gtcgtggtat ccatgctttt tttcaactaa taacatcatc tctcttcatg accagt 2216

<210> 78
<211> 5320
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2804937CB1

<400> 78
tcgccgtgaa cgcagcacat ttaggcgcca ggcagtacgg cgccggcaca atgcagggag 60
taaccctacc cctcctacat tgctcatcgg atcacccta agccttcaag atggtcagca 120
aggccagcag tccacagccc aggtcaaagt ccagtcccgc cccccttccc aggctgcagt 180
gctcagtgct agtgcctcct tgctggtgag aaatgggagt gtccacttag aagcatcaca 240
tgacaatgca tctgctgtag gcggtagcag tttgcacgat gaacttggta agttctcttc 300
tacgctgtat gagactggtg gctgtgatat gtcacttgtg aattttgaac cagcagcaag 360
aagagcatcc aatatctgtg acacagattc tcatgtatcc agttctacct cagttcgatt 420
ttatccacat gatgtgctct ctctcccaca gattcgattg aatagactat tgaccattga 480
tacagatttg ttggagcaac aggacattga tctaagccct gacttggcag ctacttacgg 540
cccaacagaa gaagctgccc aaaaggttaa acactattat cgcttttgga tcctacccca 600
gctgtggatt ggcattaact ttgacagact cacacttttg gccctgtttg atagaaatcg 660
tgagatcctg gaaaatgtgt tagctgtcat cctggctatt ctcgtggcct ttttgggatc 720
tattcttctc atacaaggat tcttcagaga tatctgggtc ttccagttct gcctcgtcat 780
agccagctgt caatactcac tgcttaagag tgttcaacca gattcttctt ctcccagaca 840
tggtcataat cgtatcattg cctacagtag accagtttat ttctgcatat gttgcggtct 900
tatttggctc ttggattatg gtagcagaaa cctgactgca accaagttca aattatatgg 960
aataactttc accaatccac tggtgtttat atcagccagg gatttagtta tagtgtttac 1020
actctgtttc ccaatagtgt ttttcattgg tctcctgcct caggtgaata catttgtaat 1080
gtacctttgt gaacaattgg atattcatat ttttggtggt aatgccacta caagcctgct 1140
tgcagcactt tacagtttta tctgtagcat tgttgcagta gccttattgt atggattatg 1200
ttatggggct ttaaaggatt cttgggatgg ccagcatatt ccagtacttt tctccatttt 1260
ttgtggttta ttagtggcag tgtcttacca tctcagccga caaagcagtg atccatctgt 1320
acttttctct ttagtgcaat ccaagatttt tccaaaaacg gaagagaaaa atccagaaga 1380

ccctctatct gaagtaaaag atccactgcc tgaaaaactt agaaattctg ttagtgagcg 1440
attacagtct gacctggtag tatgcattgt aattggtgtg ctgtattttg ctattcatgt 1500
aagcacagtc ttcacagtat tgcagcctgc cctcaagtat gtgttgtata cattggttgg 1560
ctttgtgggt tttgtaaccc attatgtgct gcctcaagtt agaaaacagc taccatggca 1620
ctgtttctct catcctctgc taaagacact agagtataat cagtatgaag ttcgaaatgc 1680
agccactatg atgtggtttg agaaacttca tgtgtggctt ctttttgtgg agaagaatat 1740
aatctatcca ttgattgttc tcaatgaact gagcagcagt gcagagacaa ttgctagtcc 1800
aaagaaactg aatacagaat taggtgcttt aatgatcact gttgctggtt tgaagttgct 1860
acgatcctct tttagcagcc ctacatatca gtatgttaca gtcatcttta ctgtgctgtt 1920
tttcaaattt gactatgaag ctttttcaga gaccatgctg ttggatctct tctttatgtc 1980
catactcttc aacaagcttt gggaactact ttataaattg cagtttgtgt atacctatat 2040
tgccccatgg cagatcacat ggggttctgc tttccatgct tttgctcagc cttttgcagt 2100
gcctcattca gccatgctgt ttattcaggc tgctgtctcg gccttcttct ctactccact 2160
gaacccettt ctgggaagtg caatattcat cacttcatat gtccgacctg tgaaattctg 2220
ggagagagac tataacacaa aacgagtgga tcattcaaat accagattgg cttcccagct 2280
tgatagaaat ccaggatcag atgacaacaa tctgaattcc atcttttatg agcatttaac 2340
tagatcccta cagcacagcc tctgtggtga tttgctacta ggacggtggg gaaactacag 2400
tacaggggac tgtttcatcc ttgcctctga ctatctcaat gcattagtac accttataga 2460
gataggcaat ggtctggtca cttttcagct gcggggactt gaattcagag gtacctactg 2520
tcaacaacgg gaagtggagg ccattactga aggtgtagag gaagatgaag gattttgctg 2580
ttgtgaacct ggccatattc ctcacatgct ttcatttaat gctgcattta gccagcgatg 2640
gctagcttgg gaagtgatag tcacaaagta cattctggag ggttatagca tcactgataa 2700
cagtgctgct tctatgcttc aagtctttga tcttcggaaa gtactcacca cttactatgt 2760
caagggtatc atttattatg ttacgacctc gtctaagcta gaggagtggc tagctaatga 2820
gacaatgcag gaaggacttc gtctgtgtgc tgatcgcaat tatgtcgatg tggacccgac 2880
ctttaatcca aacattgatg aagactatga ccaccgactg gcaggcatat ctagggagag 2940
tttctgtgtg atttacctca actggataga gtactgctct tcccgaagag caaagcctgt 3000
ggatgtggac aaagattcat ccctagtgac tctctgttat ggactctgtg ttctgggacg 3060
gagagctttg ggactgcat cccatcatat gtccagtaat ttagagtcat tcctctatgg 3120
attgcatgcc ctatttaaag gagatttccg tatttcttca attcgagatg aatggatctt 3180
tgctgacatg gaattgctaa gaaaagtagt agtccctggg atccgtatgt ccattaaact 3240
tcatcaggat cattttactt ctccagatga atatgatgac cctactgtgc tctatgaagc 3300
catagtatct catgagaaga acctcgtaat agcccatgaa ggggaccctg catggcggag 3360
tgcagtactt gccaactctc cctccttgct tgctctgcgg catgtcatgg atgatggcac 3420
caatgaatat aaaatcatca tgctcaacag acgctacctg agcttcaggg tcattaaagt 3480
gaataaggaa tgtgtccgag gtctttgggc agggcaacag caggagcttg tttttctacg 3540
taaccgtaac ccagagagag gtagcatcca aaatgcaaag caagccctga gaaacatgat 3600
aaactcatct tgtgatcaac ctattggcta cccaatcttt gtctcacccc tgacaacttc 3660
ttactctgac agccacgaac agcttaaaga cattcttggg ggtcctatca gcttgggaaa 3720
tatcaggaac ttcatagtgt caacctggca caggcttagg aaaggttgcg gagctggatg 3780
taacagtggt ggcaatattg aagattctga tactggaggt gggacttcct gcactggtaa 3840
caatgcaaca actgccaaca atccccacag caacgtgacc cagggaagca ttggaaatcc 3900
tgggcaggga tcaggaactg gactccaccc acctgtcaca tcttatcctc caacactagg 3960
cactagccac agctctcact ctgtgcagtc gggcctggtc agacagtctc ctgcccgggc 4020
ctcagtagcc agccagtctt cctactgcta tagcagccgg cattcatccc tccggatgtc 4080
caccactggg tttgtgcctt gtcggcgctc ttctactagt cagatatcgc ttcgaaactt 4140
gccatcatcc atccaatccc gactgtcgat ggtgaaccaa atggaaccct caggtcagag 4200
cggcctggcc tgtgtgcagc acggcctgcc ttcctccagc agctccagcc aaagcatccc 4260
agcctgcaaa catcacactc tcgtgggctt tcttgcgaca gagggaggtc agagcagtgc 4320
cactgatgca cagccaggca acaccttaag tcctgccaac aattcacact ccagaaaggc 4380
agaagtgatt tacagagtcc aaattgtgga tcccagtcaa attctggaag ggatcaacct 4440
gtctaaaagg aaagagctac agtggcctga tgaaggaatc cggttaaaag ctgggagaaa 4500
tagctggaaa gactggagtc cgcaggaggg catggaaggc catgtgattc accgatgggt 4560
gccttgcagc agagatccag gtaccagatc ccacatcgac aaggcagtgc ttctggtcca 4620
gattgatgat aaatatgtga ctgtaattga aactggggta ctagaacttg ggctgaagt 4680
gtgagccagt gtttattata aagacatttc tttttccctc tcaattccaa ggcattggaa 4740
aaagagagga acaagcagaa gatgcctgca ggtatcactt tgatcctatg tgggagcgac 4800
tgaaaataga atgagcttgg ttaagcacct ctcctttgcc cttcaccctg actcctgtca 4860
ctgtctccat ccccaaataa agctgaaata ttttttttaag ttagctgccg agaaaacatt 4920
ttgcatgaag gataaagttc tgttaaaata catccttaaa aaaagttttt cctatgcatt 4980
gcctatgctt taaatcaaca aactcttcat ttctaaacta tgatctcata tttttctaat 5040
ttctttgcca aaaataattg ccatgttttg tcatagaaca tgaaattcat ttaccttgat 5100
ttttaagaac agaccagtgt agaaatgttc cttttgactt atagcagtat aaaagtttaa 5160

tgtacagtga aagagactat gaacagacat agatttatct tattccttga gcgctaaaaa 5220
ctttaataaa aaaaactgca ctaatttttt acagcaatgc actaaagtct gagatttctc 5280
agaacattta tttatatata aaaataaaat accattattt 5320

<210> 79
<211> 653
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7494915CB1

<400> 79
tgcctttcat gtgaagaacc aggggctact ccaggctaga tttccccatg atgccaggag 60
agaaaaaaca gtcaggctcc cagcaacaga acaatgggca ggcaagcaaa aaccgcacac 120
agaaagaaat agtaacccag aaaagaccaa taacaagcaa tgagattgaa ttagtagtta 180
aaaaaaaatt gccaagagaa aaaggtccag gaccagatgg attcatagct gaattcttcc 240
ggacagtcaa agaagaattg gaaccaaccc tactgaaact attccaaaag attgagagag 300
aaagaatcct ccctaacaca ttttatggag tcagtatcac ccttatgcca aagccagaaa 360
aggacacaac agcaacaaca acaacaacaa caacaaacta cagaccaaca tccctgatga 420
atgtagattc aaaaatcctc aacaaaatac tagctaacca aatccaacct cacatcaaaa 480
agataataca ccataatcaa aagctcttta gtttaattag atcccatttg tcaattttgg 540
cttttgttgc cattgctttt ggtgttttag acgtgaagcc tttgcccata cctatgcact 600
gaatggtaat gcctaggttt tcttctaggg tttttatggt tttaggtctt aca 653

<210> 80
<211> 1794
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2073751CB1

<400> 80
aattcggctc gagaggagtg gtccggggca ccggtttcaa agtcgcaggg cgggccgagt 60
ggacttccgc tgctggcctg gggcttccca gccgtcttgg cgttgtcctc tccaaccccc 120
gccgctccgc gtagaacgcc gctctcaggc tgccgtcaag ctcccgcggc actctcctag 180
gtggcccgac gagacccaga gtgacccgcg ggacgcctgt atcgaccgcg tcctcttccc 240
accagcgtgg gattcggttg aacgtggagt ccccagcaat cttcagtctc tcaccagggc 300
cagggactcg tctggggcgc gggggaaaga agcgtggcgg ggctgtagat gccgcgtgag 360
taggatgcag attgcaccgc tggagcgctt gacaaccaac cgagcgttgg cttagttttg 420
ttttcccgca cagcaagctc tctgtctttc agaggaaggt aaaggtggtg aaagctctaa 480
actaaaattg tatcgaaatg gctgcagaaa tcgactttct gagagagcaa aatcgaagac 540
taaatgaaga ttttaggcgg tatcaaatgg aaagtttttc caaatattca tctgtacaga 600
aagctgtctg ccaaggagaa ggagacgaca catttgaaaa cctagtattt gaccaaagct 660
ttttagctcc tcttgttact gagtatgata aacacctagg agaactaaat gggcagctga 720
aatattacca gaaacaggtg ggtgagatga aattacaatt tgaaaatgtc atcaaggaaa 780
atgaaaggtt gcacagtgaa ttaaaagatg ctgttgaaaa aaaattggag gcctttcccc 840
tgggcacaga ggtaggaact gacatatatg cagatgatga aacagtcaga aaccttcaag 900
aacaattgca gctagccaat caagaaaaaa ctcaggctgt ggaactctgg cagactgttt 960
ctcaggagtt ggacagacta cacaagcttt accaggaaca tatgactgag gcccagattc 1020
atgtatttga aagtcaaaaa caaaaggatc agctatttga ttttcaacaa ctgaccaaac 1080
aacttcatgt tactaatgag aacatggaag tgactaacca acagtttctg aaaacagtaa 1140
ctgaacaaag tgtgataatc gaacaactcc gaaaaaaact taggcaagcc aaattagagc 1200
tgagagttgc tgtagcaaaa gtggaagagt taactaatgt gactgaagat ctgcagggac 1260
agatgaaaaa gaaggagaag gatgtggtgt ctgcccatgg aagagaggaa gcatcagata 1320
ggcgtttaca gcagttacag tctagtataa aacaattaga aataagatta tgtgtgacaa 1380
tccaagaagc caaccaatta agaaccgaaa atacacatct ggaaaaacag accagagagc 1440
tacaagcaaa gtgcaatgaa ttagaaaatg agagatatga ggctattgta agagccagaa 1500
atagcatgca actcttagaa gaagctaacc ttcaaaaaag tcaggctcta cttgaggaga 1560
agcaaaaaga agaagacata gagaaaatga aagagacagt ttctcggttt gtacaagatg 1620

ctaccataag aaccaagaaa gaagttgcaa acaccaaaaa acaatgtaat atacaaattt 1680
ctcgattaac agaagaactt tcagcccttc aaatggagtg agctgagtcg acagaaactt 1740
catacccaag agctgctttc tcagctggaa atggcaaatg aaaaggtagc tgag 1794

<210> 81
<211> 2691
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3178841CB1

<400> 81
gcgaccagcg gaaggaggga gggggccgcg ctcggcgccc cggccgggcc actgggccac 60
aggccacgcg gccacgcagt ccgagcggga gccgagccgg gcggggcgag ggcagctccg 120
ggaaggaacg tcccagggat ggaagtgctt ggatgcggtg ctgctggctg cggatgtgcg 180
caaggagatg ggatggagag cctgagttgg cattcgtata aatgacctgc ctggctccca 240
ccatgagtgc tgagcttaac gtgcctatcg acccctctgc tcctgcctgc cctgagcccg 300
gccataaggg catggattac cgggactggg tccgccgcag ctacctggaa ctggtcacct 360
ctaaccacca ctcggtacag gccctgtcgt ggcggaagct ctacctgagc agggccaagc 420
tgaaggcctc cagcaggacc tccgccctcc tctccggctt tgccatggtg gccatggtgg 480
aggtgcagct ggagacgcag taccagtacc cgcggccgct gctgattgcc ttcagcgcct 540
gcaccacggt gctggtggcc gtgcacctgt tcgccctcct catcagcacc tgcatcctgc 600
ccaatgtgga ggccgtgagc aacatccaca acctgaactc catcagcgag tccccgcatg 660
agcgcatgca cccctacatc gagctggcct ggggcttctc caccgtgctt ggcatcctac 720
tcttcctggc cgaggtggtg ctgctctgct ggatcaagtt cctccccgtg gatgcccggc 780
gccagcctgg ccccccacct ggccctggga gtcacacggg ctggcaggcc gccctggtgt 840
ccaccatcat catggtgccc gtgggcctca tcttcgtggt cttcaccatc cacttctacc 900
gctccctggt gcgccacaaa acggagcgcc acaaccgcga gatcgaggag ctccacaagc 960
tcaaggtcca gctggacggg catgagcgca gcctgcaggt cttgtgaggg gccgagggcc 1020
ggggctggga gcggccctgt gcccgggagt ccgcagaggc ggggatttgt cagatgcaga 1080
cattttgcaa ggctgccggg tagttcaaga ccaaagtttt cctcttgtct taataccata 1140
aggactggat gacttctcct gagatagaac cgtttggttc aatgagggac tgtgttgcta 1200
agagcgttgg gggcaaagcc aggctggttc cttggcctcg gggtttcctg ggtcggggac 1260
acggtgaaga ggctccagcg ggacctgccc atcagtcctg ggccaggagg ggctccaagc 1320
agcacccagc ggtccggggg agtctcagac ccggcatgcg tggctggcag acctgggaga 1380
gccagggcag ggttttgcgt tcagagaagg attgccccag agacccgtgg tggacttcat 1440
gggtgctgag tggcccgtgt gacagtgatg acacgaaggc ttcggcgttt gagtgggtgc 1500
aggtgcacgc cagggcttgg tgcttccctg cctggccctg gagggagctg ggtggcctgg 1560
cttcagggga agacaggagc caggacacac gtcagcccag caggtgtggg gggtgctgca 1620
gccctcggca gtggggtcag gccctggggg atgtttccaa tggtgggcag cctggccagg 1680
ccggagaaga catgttcacg ggcatctatc agatgccccc ttgaggaggc tgagttattt 1740
gagggctgct gcaaagtacg ctaggctcaa attctctttt cccagccaga gccctggcca 1800
cacggactca gaggggccac cggggtgggg aaaggacccc tccccgaccc cccgcagcca 1860
ctggcctcca gctctcggcc acagaatggc ctctaaggct gactcagccg ctccccttggg 1920
ctgtggcagc aggaggcggg ggctctggct caggccccgg agcctgtgca gcttgcccat 1980
ggccctaggc agcgagggga cagcctgggg gacttcctgc ctaggcaagg tcattggccg 2040
ggcctggcct gtggatagtg gggccagggg ccggcccagg ccaaatgagt gccctccttg 2100
ttatgacacc aagtgactac aagggaggca agacccctcc aggcctctca gccgacactg 2160
ggtcccacca cacacagtga ctgtgccgtg cagtgcaggt tctggccttt tccttgaagg 2220
catctggtag acccgaagcc acgctctcgg gccgcacatg cacgccgcag caccagctgc 2280
cctgagctgc ttgtacaacc aaacaccttt cccctcttct ccagctgtaa cctggagagt 2340
cagccatgcc ttgtcttttg ttctcataaa tagtcactgg ggccgggcgc agtgactcac 2400
gcctgtaatc ccagcacttt gggaggccta ggtgggcgga tcacttgagg tcaggagttc 2460
gagaccagcc tggccaacat ggtgaaaccc tgtctctact aaaaaaatac agaaaattag 2520
ctgggcgtgg tggcgggcgc ctgtagcccc agctacttgg gaggctgagg cgggagaatg 2580
gcaatggcgt gaacccggga ggcagagctt gcagtgagct gagatggcgc cactgcactc 2640
cagcctgggc gacagagcca gactcaatct caaaaaaaaa aaaaaaaaaa a 2691

<210> 82
<211> 2056
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3674807CB1

<400> 82
gcagctgata gaaccatggc gaccattgct gctgctgcgt ttgaggccct catggatgga 60
gtgacatgct gggatgtccc cagaggcccc atccccagtg aactccttct tattggagaa 120
gccgccttcc ccgtgatggt gaatgacaag ggccaggtgc tcattgctgc ctcctcctac 180
ggccgaggcc gcctcgtggt tgtgtcccat gagggctacc tgtcgcatgc tggcttggct 240
ccatttctcc tcaatgcagt gagctggctc tgtccctgtc ctggggctcc cgtgggagtg 300
catccatccc tggcacctct agtaaacatc ctacaggatg ctgggcttga ggcacaggtc 360
aagccagaac caggagagcc cctaggggtt tactgtatca atgcctacaa tgacaccttg 420
actgcaacgc tgatccagtt tgtgaaacat ggagggggct tgttaatcgg gggccaggcc 480
tggtactggg ccagccagca cggccctgac aaggtgctct ccaggttccc tgggaacaag 540
gtgacaagtg tagccggagt gtacttcact gacacctatg ggacagaga ccggttcaag 600
gtctctaaga aggtgcccaa gatcccactc catgtcagga ggtagagccc ggtactcacc 660
cagcaatcat tgttgcctga caaaagtcac agagaaggct cagttcttcc ttcctatgtg 720
agtatatctc cacttggggc ctcccagccc ctcccagtgg tacagcttct cccctgtgta 780
tctcaagtca tccatacttc ttagactcat cccctctaaa ccttaagcag tgtcccaatc 840
cccatcttct tgaaagtacc tgatttttgt atttttagta gagacggggt ttcgccgtgc 900
tggccgggat ggtctggatc tcttgacctt gtgatccgcc agcctcagcc tcccaaagtg 960
ctgggactac aggtgtaagc caccatgccc ggcctggtct ttcttagtga tacataaaat 1020
aatggcattt ctcacaatca atttgtcata gatttgagga aacgctatag aggtagtgcc 1080
cttacttgca tatcttaaaa tccacatttt ctaaatgaac acctccagga attcctccca 1140
tttagtcaac tatgtgttga atattttggc ttcagtttct tgactttaa aataattaaa 1200
catataaaca agttagaaaa actagtttat tatgcccaga gaaagtagta ctgcttatga 1260
atttctcgat ctaactttct gcaaaggaag gctgtaaaaa caccaaatct ggactctccg 1320
tagggtagtt ctgacctagg aacacatcac aattaaattg tttaacacta cgcttgaaga 1380
tttatacaaa cacacactgt ttttaaaact tcacttagat ctggttagag tattcacttc 1440
ttcttaagga atggtcaaaa tgtaatgaat caagaaaagg ctttgtattt caagaatttt 1500
tttaaaacaa tttgtatata ctgtaaatgt caagcaatag gaggttggtt aaataaatta 1560
tggttaacaa tagattgcaa accatatagt gattaaaata attattaggt aacttgaaaa 1620
atgcttaatg gaaaaagttt taaaattaga ttccaggata ccatgctctg aataagccaa 1680
attttggggt aaaaaaaata aaaaaccatg aaaccagaca tagaagcaga acctaaagca 1740
tatttatcaa aatatctttg tttaaaaaaa atgtcttcag tggtcgtctt ggctggtggg 1800
attttaggtc attttaatta tctttataaa actttattct ctcgaatttt aaaattaaca 1860
tctatttctc ttataattaa aaattaaatg ttatattttt taaaaagtat agtttattta 1920
ttccagtagt aaaacattgt tgaaaaatgt aactcttgtg cagattctct ttaatttctt 1980
ttcagcatga aaatgaggag atggcatgtc ttagttttgt catcccattt tcagactcac 2040
atggcaaatg tcatag 2056

<210> 83
<211> 2275
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1794922CB1

<400> 83
ctttcatatc cttagactgt atctgaaatt ttcatagctt tgcccacctt ttagggttca 60
gagttactac ccattcagtg tcggtggcct ttagctgttg gcaagcttcc aagtaaatag 120
ctcttttgca gagatctcta cccattctat tcttccactc ctgggtagtg gagaaccaaa 180
gttcacctga gtaatgcccg tgcccagcac ccagtttcct gcaggggtta gcacggatat 240
tggaggctcc cccgatccca cagatactca ggatgggtat tggagggggc ttggttttta 300
agaggacatg tgaaagtaat taacccgatt tacatttgta taatgagtcc aaacaccacc 360
ttgcaaaaga actgagagtt tagcctgctt aatggagttt ctgcttttca cggtccttgg 420
ccttaccgac ggtttacagt ccagccagat ttagttcaca aaaaaccaaa accaggtcga 480
cctcactgca gtatatgatg gcagcgagcg ttttaaagcg caggaccttt agtctcctca 540
atagttttgt gtaaagttgt ccaaaagtgt tttttttaa agtgagatga tccagtcggt 600

aaacgctttg aaaaagaata gctgtcccct gagcgggcat tagtcgcgtg tgaggtcagg 660
gcccatcatt gcgcgtgttg ggagcgccgc gctggtctat gagcgagcgc ctcggcttgt 720
ggctgggccg ggccggggcg gggctgtctt cccgcggagc gtgctggggg cgggtcgcgc 780
cgggccgggc cgctagtgcg catgggcggg cgtcctcggc tctaactgcc gccactttcc 840
acacgctggg agggccgtta cctcagagat acccgtggcc ggcatgttgg ttgaaaaagc 900
ttcccggaag ggagacgaag agaaaggaga ggagcagctc gtgatcatcc ccggtagcga 960
gtacgcggcg aagtaggcgg cggcggaggg agcgctgatg aagatggatg tgtcagtgag 1020
ggccgcgggc tgctccgacg acctcagctc tggggaggcc gacgtagacc caaagctcct 1080
ggagctcacc gctgacgagg agaagtgccg cagcatccgc aggcagtacc ggcagctcat 1140
gtactgcgtg cggcagaacc gggaggacat cgtgagctcg gcgaacaact ccttaaccga 1200
ggctctggag gaagccaacg tcctctttga tggcgtgagc cgaaccagag aagcagccct 1260
cgacgcccgg tttcttgtta tggcttctga tttgggtaaa gaaaaggcaa agcagttaaa 1320
ctcagatatg aacttcttta atcagttagc attttgtgac tttctgtttc tgttcgtggg 1380
tctgaattgg atggaaggcg atcctgacaa gttgagtgat tgtgatgata gcatagctct 1440
ttccttctgg aaggcaatag aaaaggaagc aacatcctgg atggtaaaag ctgagacatt 1500
ccattttgtt tttggttcat tcaagctaga acgttctgca ccaaagcccc gacttgaaca 1560
ccagaaaaaa gttcgcaaga tggaagaaaa tggcaacatg cctacaaagt tgcagaagtt 1620
ggacctgagt agttatccag aagcgacaga aaaaaacgta gaaaggattt tgggattgtt 1680
gcaaacctac tttcgaaagt atcctgatac tcctgtgtcc tattttgagt ttgtgattga 1740
tccaaactct ttttctcgta ctgtggagaa tatattttat gtttctttta ttgtaagaga 1800
tggttttgca agaataaggc ttgatgaaga caggctgcca atattagagc cgatgaatgt 1860
taaccaaatg ggtgagggaa atgattccag ttgccatggc aggaaacagg gagttatatc 1920
tttgacttta caggagtgga aaaacattgt ggcagctttt gaaatttctg aggctatgat 1980
tacatactcc tcatactaaa gatttcttag tatagcatcc tttttgtgtt ttttttctga 2040
agttagatgg agagtaaaat gtaaactgaa gcacatattg tatctcttgt aaagtgaaaa 2100
agtattttca agaacatcag acattgtttt actgtgcagc atatttttct tagtaattta 2160
taaggtcatg atcttctgtt attaaaacaa attcactggc atatttatgg gaacgtttta 2220
ttgaatgccc tttaagacta ttaataaaaa caagttttgg ataccaaaaa aaaaa 2275

<210> 84
<211> 1219
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1795509CB1

<400> 84
gcttgttgca gctctagccc aggtcctgcc ctcttcccgc cccgccccta gggtccagct 60
cccttcacct aggagctgcc aaacatctgg atcaacctgg gcactacgag ggttgaatt 120
tctaccatta tcgcgccttt tgatattttt ttccagacct cctgctcaca tccgtaaagc 180
ccactgattc ttttactaca cttttatga gaacaagaca ttttctagga agatggtggc 240
agaaaaagag accctgagct taaacaaatg cccagacaag atgccgaaga ggaccaagct 300
gctggcacaa cagccgctcc cggtgcacca gcctcactct ctggtttctg agggtttcac 360
agtcaaagcc atgatgaaaa actcagtcgt gagaggccct ccagctgcag ggcatttaa 420
agaaagacca accaagccca cagcatttcg aaaattctat gagcgaggtg acttcccaat 480
tgcccttgag catgattcga aaggaaacaa aatcgcctgg aaggtagaaa ttgagaagct 540
ggattaccat cattatctgc ctctgttttt tgatgggctt tgtgaaatga catttcccta 600
tgagtttttt gctcggcaag gaatccacga catgctggaa cacggtggga acaagatcct 660
acctgtcctt ccacagctca ttatcccgat aaaaaatgcc ttgaacctcc gaaaccgaca 720
ggtcatctgt gtcactctca aggtcctcca gcatctggtt gtgtcagctg agatggtggg 780
caaggccttg gtgccttatt accgtcaaat cctccctgtc ctgaacatct ttaagaatat 840
gaatgtgaac tccggagacg gcattgacta cagccagcag aagagggaga acattgggga 900
cttgatccag gagacactgg aggccttcga gcgctacgga ggagaaaatg cctttatcaa 960
cattaagtac gtggtcccaa cctacgagtc ttgcttgcta aactaacagt ggcagcagct 1020
gggacttgaa acctcccgtt ggtgttggga tcatctgtct ctgttgcttt tagcatctca 1080
ttccttgtga cttccacagc tttcttttct acagctgcta aaatagtggc ttatgggcca 1140
ttggactgtt agccctattg agagcaaggc tttccaatac ataaatagtg cctgtttctt 1200
agattaaaaa aaaaaaaaa 1219

<210> 85
<211> 1015

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2017180CB1

<400> 85
ggaagcgggg gtgcagcgcg gcagaatgag ggttgattcc tcggctgacc ccacaatgtc 60
gcaggagcaa gggccggggt cctccacgcc tcccagttct ccgacacttc ttgacgctct 120
gctccagaac ctttacgact ttggaggtac agaaggtgaa acagaacaga agaagatcat 180
aaagaaaagg gaaaacaaga agagagatgt gatggcttca gcggccttgg cagcagagcc 240
atctccccta cctggttctc tcataagagg ccagaggaag agcgcttcga gcttcttcaa 300
ggaacttaga gaagagcggc attgtgctcc ttctgggacc cccacaggac cagagatcct 360
tgctgctgca gttcctccct cttccctaaa gaacaatagg gaacaagtag aagtggtaga 420
atttcacagc aataaaaaaa gaaaattgac gccagatcat aacaagaaca caaagcaggc 480
taatcctagt gttttggaga gagatgtgga tacacaagaa tttaacctag aaaaagctcg 540
tttagaagtg caccggtttg gtatcacggg ttatggaaaa ggaaaggaga gaatcctgga 600
acaggaacgt gccattatgc tgggcgctaa gcctcctaaa aagagttatg tgaattacaa 660
ggttttacag gagcaaatta aagaaaaaaa ggcagcaaag gaagaagaaa agagactggc 720
ccaagaaaca gatattttca agaaaaagaa gaggaaagga caggaggaca ggaaatccaa 780
aaagaagtcc gctcccagta ttttgtcaaa tggacggatt ggacaggttg gaaaattcaa 840
aaatggaaca ctgattctga gcccagttga tatcaagaaa ataaattctt ccagagtggc 900
caaatgaagt actttgtcaa ataaaacgag gatgggaaca actgagtcaa ctgcagcggc 960
cctggaccct gtgcgactct agatagatcg ttcttaacac gcatttcaca gactg 1015

<210> 86
<211> 2392
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 219442CB1

<400> 86
ggcggagtga ccacatagag taaaaagaat gtaaccctct ttggttttat gtgttttggc 60
tcaggttgga catacttata taaatgtgat tgacattgaa gctaatgatc ttctacagga 120
attacctgtg agagaagagc cttcaaatga taatgttatc aaacagcaaa gcgatcatct 180
agcagttcca tcgtctgcag agttacatta tatggcagct tcagttacta atgctgttcc 240
cccacataat tttaagagtc aagaagtaac tccagcttgt ctggatggaa aaagcttgag 300
agcaggcatt acagaagtga aggagcctag tgtcacctca cctacaccat cagacataca 360
gcagaacaaa ggtctgccaa aaccagagtt ccgattcaaa ggacagagca caaagtcaga 420
ctctgcagaa gattatctat tgtggaaacg gctgcaaggt gtctctgcag cttgccctgc 480
accaagctct gcagctcacc aactagagca tctcagtgct aagcttcaga aaattgacga 540
gcagttgcta gcaatacaga acattgctga aaacatagaa caggatttcc ccaagcctga 600
aatgctagat ctacattgtg ataagattgg accagtggat cacattgaat tctcttctgg 660
ccctgaattc aaaaaaacat tagcttcaaa aaccattagc atttctgaag aagtgcgttt 720
tttgacccat atggatgaag aagatcaaag tgacaaaaag gagacttcag aacctgaatt 780
ttcaataaca gaaaattatt ctggtcagaa aacctgtgtg tttcctactg ccgattcagc 840
tgtcagcctt tccagttcca gtgatcagaa tactacttct cctggtatga atagcagtga 900
tgaattgtgt gagagtgttt cagtacatcc gctccagatg actggattga ctgatattgc 960
agacattatt gatgacctta taattaaaga cggagtttcc agtgaagaac ttggcttaac 1020
agaacaagct atgggcacct ccagaattca gcattattct ggcagacatt cacaaagaac 1080
tgacaaggaa agaagagaga ttcaagcctg gatgaaaaga aaacgaaaag aaagaatggc 1140
aaagtactta aatgagctgg cagaaaagag agggcaagaa catgatcctt tctgtcccag 1200
aagcaatcca ctttacatga cttcaaggga aataaggctg agacaaaaga tgaagcatga 1260
aaaagacaga ttgctgctct ctgaacacta tagtcgtcga atctcacaag cgtacggtct 1320
gatgaatgaa ctgttatctg agtcagtaca gctaccaact ctaccacaga aaccattgcc 1380
taacaaaccc agccctactc agtcttccag ttgtcaacac tgcccttctc caagaggaga 1440
gaatcaacat ggtcacagtt ttctaataaa tcgacctgga aaagtcaaat atatgtccaa 1500
accgagttat atccataaga ggaagtcttt tgggcaacct caaggctcac cttggccaca 1560
tggaactgcc actttcacca tacagaaaaa agctggtgga gccaaagcag cagtaagaaa 1620

ggctacgcag tctccagtta ccttccaaaa aggctctaat gctccgtgtc atagtctgca 1680
gcatacaaaa aaacatggaa gtgctgggct tgcacctcaa accaagcagg tgtgtgtaga 1740
gtatgaaaga gaggagactg tggtgagtcc ctggacgata ccttcagaaa tccataagat 1800
tcttcatgag agtcacaatt cccttctaca agacttgtct ccaactgaag aggaagagcc 1860
agagcatcct tttggggtgg gcggtgtgga cagcgtgtct gagagcactg gcagcatcct 1920
cagcaagctg gactggaatg ccatcgaaga catggtggcc agcgtggagg accagggcct 1980
gtctgtccac tgggccctgg acctgtaaga cctggatatc attgggtttc catgcacagg 2040
ccagcacctc agtaatgtgg ttctgaaaga ttaacaggtt taagggacag aagcaatgaa 2100
agaagcaatg tgaattttcc atttgctttc atattattac ctggattagc cattaccaga 2160
ggaaaaataa acatttctca gtaactttgc ctttatgggg aaagggttga ctattgatgt 2220
attatatgtt tttgtatttg atgcatcatt aggcataatt tttaaaatga taagtacctt 2280
tcaagccaag tttgcataac ctactttcaa taaaaaccct ctatcttgcc tcctccttta 2340
ttaccctctg agttttgaga aacaaccata tacagatgaa tctaatagga aa 2392

<210> 87
<211> 1799
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2597459CB1

<400> 87
gcagagcgaa gctactgcgg gttctgttaa cctcagcatc gtggggcgaa gcagagccat 60
tgtgcatcaa ggagaggccg gtgcctgcgc tgccgtctct ggcacctaac ccagcagacc 120
gctcacccca tcggctggaa tgcttgctga ttcctctgca cacttctaag gctgagaacc 180
tgaggaaccc agctggaaaa tgccgtctga acgctgcctc agtattcaag aaatgctgac 240
aggccagagg ctctgccact ccgaatctca caatgacagt gtcctggcag cgctgaatca 300
gcagaggagt gatggcatcc tctgcgacat caccctgatt gctgaggaac agaaattcca 360
tgctcacaag gcagtcctag cagcatgcag tgactatttc cgggcaatgt tcagtctttg 420
tatggtggaa agtggagctg atgaggttaa tttgcacggt gtgaccagcc ttggcttaaa 480
gcaggctctg gagtttgcat acacaggaca gattttgctg gagccaggtg tgatccagga 540
tgtgctagca gcgggcagtc acctacagct gttggagctt ctcaatttat gctcccacta 600
tctcatccag gaattaaata gctttaatta cttggatctg tacagacttg ctgacctctt 660
taacctcact ttgttggaga aggcagtgat cgatttctta gtgaaacatc tctctgaact 720
cctgaagagc cgcccagaag aagttctaac gcttccctat tgcctgcttc aggaggtgct 780
gaagagcgac cgcctgacct ccctgagtga agagcagatc tggcagctag ctgtgaggtg 840
gttggaacac aactgccact accagtacat ggacgagctc ctgcaataca tccgctttgg 900
cctaatggat gtggatactc tccatacagt tgccctgtcc cacccccttg tccaagcaag 960
tgagactgca acagcccttg tcaacgaggc cctggaatac caccagagca tctatgcaca 1020
gcctgtctgg cagactcgca ggaccaaacc acgattccag tcagacactc tgtatatcat 1080
tggtgggaaa aagcgcgagg tctgcaaggt caaggaactt cggtacttca atcctgttga 1140
tcaggagaat gctctcatag ctgccattgc caactggagt gagctggctc ccatgcctgt 1200
gggaaggagc caccattgtg tggcagtcat gggggacttc ctgtttgtgg caggagggga 1260
agttgagcat gccagtggcc ggacgtgtgc tgtgaggact gcctgtcgct atgacccccg 1320
cagtaattcc tgggcagaga tagcacccat gaaaaactgc cgggagcatt ttgtgctggg 1380
tgccatggag gaatacctct atgcagttgg gggcagaaat gaactgcgcc aggttctgcc 1440
tacagttgag cgatattgcc ccaagaagaa caaatggact tttgttcagt cctttgacag 1500
atccctttca tgccatgctg gatatgtggc tgatggtctt ctttggatat caggtagaac 1560
atacctcatg ttggatttat caaaacacac tttcattgtg gtatatattt aaaagtcaag 1620
ctttaatact gtggccatgt gtaattgaaa gattgaacca agcatgccat tttagctaat 1680
taacattcat ttattgagca cctactctag agagcaatgt aattggaaat aatatgcctt 1740
tctgtgactt tctgagtgac taatggaata aaaattcaag gtggaaaaaa aaaaaaaaa 1799

<210> 88
<211> 3718
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2783863CB1

<400> 88
ggcgctggcg gcagtggccg gtgtgagcgc agagccgggc cctgggcagc agcggcagcg 60
cggtaggacc tcgcgcagcg tccgcgggct ccggggcggg ggcgccagcg gcgaagcccc 120
ctccccgggg aggcgggacc tgggggagct cccgagccgg ggagcggcgg cggccgggaa 180
cgatgcatca gaagctgctg aagagcgcgc attacatcga gctgggcagc taccagtact 240
ggccggtcct ggtgccccgt ggcatccgcc tgtacaccta cgagcagatc cccgggtccc 300
tcaaggacaa cccgtacatc accgacggct accgggccta cctgccgtcc aggctgtgta 360
tcaaaagttt gtttatttta tctaatgaga cagtaaacat ctggagtcat ttgctgggtt 420
tctttctctt cttcaccctg ggaatatatg acatgacatc tgtgttacct tcagcaagtg 480
cgtccagaga agattttgta atttgttcta tttgtctttt ctgcttccag gtctgtatgc 540
tttgctctgt gggctatcat ctttttttcct gccatcggtc agaaaaaaca tgtcgaagat 600
ggatggcatt agattatgca ggaatttcta ttggaatact gggctgctat gtctcaggag 660
tattttacgc attttattgt aataactact ggcgtcaggt gtacttgatc acagtgcttg 720
ctatgatcct ggcagtgttc tttgcgcaga ttcatcccaa ttacctcacg cagcaatggc 780
aaaggctccg ttctatcatc ttttgttctg tttcgggata tggagtgatt cctactcttc 840
actgggtttg gctcaatgga ggaattggtg ctcctattgt acaggacttt gcaccccgtg 900
taattgtgat gtatatgatt gctcttcttg ctttcctatt ctacatttcc aaagtcccag 960
agcggtactt tccaggacaa ctaaactacc tcggatcaag ccaccaaata tggcatatcc 1020
ttgcagtagt gatgttatat tggtggcatc agtcaacagt gtatgtcatg cagtacagac 1080
atagcaagcc ttgtcctgac tatgtttcac atttgtgaat taggtatggc cacctggtga 1140
attcagttgt taagcaatat ataatgggga attgtatacc ccactatttc taagattccc 1200
attagttttc cctttttcct ttttaatatg agtaatgctt tataaaaatg ggaaaaaaag 1260
tatacttaag gatctgtagt aataactgct ttacaaaatc cttaaaacta ctaatttgct 1320
gcttgtacag aaagtgaaaa ttagttggca atcataagaa acatctgaat aacaacgatg 1380
aatgggaaac tagtgttgaa ataggattca ttttacttag caccagctta atttccttag 1440
gaagggctca tctccattag aaaatggagt catcttatgt gcttaattat tttcagttaa 1500
ttgtcaagtt taagtgccta atcaaggcaa gtgttgtttc agcctatgct taatgcaagc 1560
taggatagtg attttaaata atcactaaaa tcactagatt taaataatca ctaaaatgat 1620
ttgtgagaaa ctggcacttc agatattata tcctttagct ataggttctt ctctccctaa 1680
gaacattaga tattttagtt ttccagaaca aaagctttaa acttctgcag taagttgaga 1740
gaagggttga gaagaggaaa agaacttctc attttctatc agataagaat cacattagaa 1800
actaagtaca agattagaca acaaattatg tggtcaaata atatagtcat tagccaccta 1860
aacattttaa ttccagatat tatttaattc catataataa ctgaattctt gtgagtggat 1920
tacaggtttt tgatcccaaa attccagagc tttcaactct ctgaatttgt agtcctgaat 1980
atcccagtgg tgggggttcc cagcattgtg ggtgctactt gcaaggccat agaatctaga 2040
tggccctgtc ttgaccctga aatgaacctt aagccttaga acaaagtcat gcagatgccc 2100
catttgataa taatcttatt cacctgtgct ctggtcctcg gtttctgcat gtgttagcat 2160
tgcattgata actcagaatc ttgataaaca cttaatattt gggcctgaag cattaaactt 2220
tctttttatt gtatatactt aaaaaataga actcactgcc ctatcataca ttgtagccct 2280
cttattcttt ggtctttcat atgcattagt taaatccctt aaagtagaca ttcataaaaa 2340
cttacattgt ttattggagt ataaaatatt acccaagttt cttcatgagt tgacatgagc 2400
tgttttaaat actggtgtat tttcagaaca gtaaaattac tgaatatcag aaaaaatgtt 2460
aattgatgat gaagcttatt cccaaaatgc cttttgtgca tatgatactt ggaaagtcac 2520
taatgtgcct cagttaatac atcagtaaaa tgttgtgttt cttttccagt gtagtgtttt 2580
tggaatataa attccccatg ctagtatagt atctcagcaa agagaatttc cccccaggag 2640
gctcagtaaa ggaataccgt gtcttaccca tcgttatgat ggaaggctgc tttgaaaatg 2700
gctgttttac cttataaggt taaaattttg atccatatgt taagtgatag aagattttgg 2760
tgcaacagta gtaggatata tttctcctag aacatccctt attggcttac atgattttat 2820
tgcctttttaa tagatatttt gtcattttgg ccaaacaaaa gacactgagt agttacactt 2880
aagttaaaaa tgaggggaaa atcattattt taggtgtgga gccattttta ttataaaact 2940
ttctcaaaat aaaaaaacat tgaatcattt caatttttgc agtccctgta ttagtatatg 3000
aatacatact tgccatttga attaataaca tgaaaagagt atactgtgtt tttaaatccg 3060
tgtttctttg aatttaaagg gtgtacaggt ctttctgtag ggaaaattat tccatgtaaa 3120
catttcaact ctgtatgaaa atgttaaata ttgtaagaaa gttatcctct cattttttca 3180
ctgctatgat atatttatta taaaataggg aatgaatgaa tgaatatgga ttgctgttaa 3240
ctagaaacac ttctgtatgt cagtcagcat ttaatgacca cctactgtgt gcacagcact 3300
actggtaaaa ttttgaagac attgttaaca ttaaaaaata ttttaaagtt gtctacaaat 3360
ctgagccttg taatgatgta tatttaagtt atttttgttt ttatagatta aagtaagatt 3420
atactatcca gttttattac taaaaaagac tggttttaat tttaccaatg tgtgaactat 3480
aaaagctttt tgcctacaga ttttacattt taaaattatc tatggctgtt ttaaattgtc 3540
tagcaattta tatggttgtg gttaactcat ttaagaaaca attatctttc tatattaagc 3600
cattttcaaa tagcaagaca gtgcttgtct ttttttgtta ttacactaac tgcaattcag 3660
taagctgcat gacaaaatat gtattatgta aataaactgg gtttactaaa tactttct 3718

<210> 89
<211> 3250
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2902971CB1

<400> 89
atggcaacaa gtatggcggc tgctagtggt agatttgaaa gtgcgaagag tatcgaagag 60
cggaaagaac agacccggaa tgccagggcc gaggtgttgc gccaggctaa agccaatttt 120
gaaaaagaag aaaggcgtaa agaacttaag cgacttcggg gtgaggatac atggatgcta 180
cctgatgtga atgagagaat tgaacagttc tcacaggaac actctgtgaa gaaaaagaag 240
aaaaaagaca agcattcaaa aaaagcaaag aaagaaaaga aaaaaaagag caagaaacag 300
aaatatgaaa aaaacaatga gtcatctgat agctcatcaa gctctgaaga tgagtgggtt 360
gaggctgttc catcccagac tcctgacaag gaaaaagcct ggaaagtgaa agatgaaaag 420
tcaggaaaag atgacaccca aattatcaag agggatgagt ggatgactgt tgattttatg 480
tctgttaaaa ctgtgtcatc atcatcactc aaagctgaaa aggaaactat gaggaaaata 540
gagcaagaga aaaaccaagc gcttgaacag tccaaactga tggaaagaga attgaatccg 600
tactggaagg atggtgggac aggtcttcca cctgaagact gtagtgtgtc atcgattact 660
aaagtttcag tggtagaaga tggtggatta agctggctaa ggaaatctta tctaagaatg 720
aaggaacaag ctgagaaaca aagtagaaac tttgaggaca ttgtagccga aagatatggg 780
tcaatggaaa tatttcagtc aaaattagaa gatgctgaaa aagctgcatc cacgaaagaa 840
gattatagac gggaacggtg gaggaaaccc acatattcag ataaagcaca aaattgtcaa 900
gaaagtagag aatcagactt agtaaaatat ggtaacagtt caagggatag atatgctaca 960
acagatactg caaaaaatag caataatgaa aaatttattg gtgatgaaaa agataagaga 1020
cctgggtctt tagaaacgtg tagaagagaa tctaacccaa ggcaaaatca agagttttct 1080
tttggcaatt tgagagctaa attcttgaga ccctctgatg atgaagaact gtcatttcac 1140
agcaagggca gaaaatttga accacttagt tcatcttcag cattggtagc tcagggctct 1200
ttgtgtagtg gttttagaaa acccaccaag aacagtgaag aaagattaac atcatggagt 1260
cgctctgatg ggagaggaga caagaaacat tcaaatcaaa agccatcgga aaccagtact 1320
gatgaatacc aacatgttcc agaagaccca agagaaaaat cacaagatga agtcttgaga 1380
gatgaccctc caaaaaaaga acatctacgg gatacaaagt ctacatttgc tggcagtcca 1440
gagcgtgagt ccattcacat cctgagtgtt gatgagaaga acaagttggg agccaagatt 1500
atcaaagcag agatgatggg gaatatggaa ttagctgaac aacttaaagt tcaacttgaa 1560
aaggcaaata aattcaaaga aactataaca cagataccaa aaaaatctgg ggtagagaat 1620
gaagaccagc aagaagtaat ccttgtcaga acagatcagt ctggaagagt atggcctgtg 1680
aacacacccg gaaaatctct ggaatcacaa ggaggaagaa gaaagagaca gatggtttca 1740
acccatgagg aaagagaaag ggtcagatac tttcatgatg atgataatct aagcctaaat 1800
gatttagtca aaaatgaaaa gatgggaaca gcagaaaatc aaaacaagct ctttatgaga 1860
atggcatcta agtttatggg aaaaacagat ggagactatt acaccctgga tgacatgttt 1920
gtctccaaag cagctgagag agaacgtctt ggtgaagagg aagagaacca aaggaaaaaa 1980
gctattgctg agcatcggag tcttgctgca caaatggaaa aatgtctgta ttgttttgac 2040
agctctcaat ttcccaagca tcttattgtt gcaataggtg ttaaggttta tttatgttta 2100
cccaacgtac ggtctcttac tgaggggcac tgcctgatag tccctttgca gcaccataga 2160
gcagctactt tgttggatga agacatctgg gaggagatcc agatgttcag aaaatcattg 2220
gtaaagatgt ttgaagataa aggattagac tgcatttttt tggaaactaa tatgagcatg 2280
aagaaacagt atcacatggt ttatgaatgt attcctcttc ccaaggaagt gggtgacatg 2340
gctcccatct attttaagaa ggccataatg gaatctgatg aagagtggtc catgaacaag 2400
aagttgatgg atctctcttc aaaagatatc agaaagtctg tacccagagg gttaccttac 2460
ttctctgtgg attttggcct tcacggaggg tttgcccatg tcattgaaga tcagcacaaa 2520
ttccctcatt actttggaaa ggaaatcata ggtgggatgc tggatataga accaagactt 2580
tggaggaaag gcatccgaga aagctttgag gatcagagga aaaaagcact gcagtttgct 2640
cagtggtgga aaccatatga cttcaccaaa agtaaaaact attgaggtgt accttccatt 2700
ttaaaatttt tcttcagatc ccgttcagtt ttatttccat tgcatctaat gaagcaactg 2760
accctcaggt cacaggcaga gagagtcaca gcagcagact gctctgggtc agtgacaact 2820
gtacatcagg agattgtgtt gttttcttct tcacctgctc ctatctatgg acttcatttt 2880
agtcaccatg gagtaaaaac atatattcga tgatctgctc acattcctgc tctctaccta 2940
cttcttggta ttccgaattg ttcagttaat ttagagaagg cctttctga agggatatct 3000
agaaagcata ggaaatcctt ggttcagagg gaagtgggtc tgcttcatgg ttgttaagtc 3060
tactgtctgg cttcttttgc atttttaaaa ataagacgat ttaattttaa tttcatagga 3120
ttttatagaa aagctcttaa ttttctacta gcagtatata gtatcacttc tattcttttt 3180

gttaaaggaa aaggaaatgt ccccaaagta ttaaagaatt aaatgatact gaaaattaaa 3240
aaaaaaaaaa 3250

<210> 90
<211> 2295
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 368660CB1

<400> 90
gtaccacccg gctcaagtag cggacacgga acagggaact atcagcccgt cggcctccgg 60
gccctgcatt ctctagccat ggaccgggac cttttgcggc agtcgctaaa ttgccacggg 120
tcgtctttgc tctctctact tcggagcgaa cagcaggaca atccacactt ccgtagcctc 180
ctggggtcgg ccgccgagcc agcccggggc ccgccgcccc agcacccgtt gcagggcaga 240
aaagagaaga gagttgacaa catcgagata cagaaattca tctccaaaaa agcggatctg 300
ctttttgcac tttcctggaa atcagatgca cctgcaactt ctgaaattaa tgaagacagt 360
gaagatcatt atgcaatcat gccaccttta gagcaattca tggagatacc tagtatggat 420
cggagagagc tgtttttccg agatattgag cgtggtgata tagtgattgg aagaattagt 480
tctattcggg aattcggttt tttcatggtg ttgatctgtt taggaagtgg tatcatgaga 540
gatatagccc acttagaaat cacagctctt tgtcccttaa gagatgtgcc ttctcacagt 600
aaccatgggg atcctttatc atattaccaa actggtgaca tcattcgagc tggaatcaag 660
gatattgaca gataccatga aaagctagca gtatctctgt atagctcttc tcttccacca 720
cacctatctg gtattaaatt aggtgtaatt agctctgaag agcttccttt atactacagg 780
agaagtgttg agctaaatag caattctttg gagtcctatg aaaatgtcat gcagagttcc 840
ttgggatttg ttaatccagg agtagttgaa ttccttctag aaaaactagg aatagatgaa 900
tctaatccac catctttaat gagaggccta caaagcaaaa atttctctga agatgatttt 960
gcttctgcat tgagaaaaaa acaatccgca tcttgggctt taaaatgtgt gaagatcgga 1020
gttgactatt ttaaagttgg acgccatgtg gatgctatga atgaatacaa taaagctttg 1080
gaaatagaca aacaaaacgt ggaagctttg gtagctcgtg gagcattata tgcgacaaaa 1140
ggaagtttga acaaagcaat agaagatttt gagcttgcat tagaaaactg tccaactcac 1200
agaaatgcaa gaaaatacct ctgccagaca cttgtagaga gaggaggaca gttagaagaa 1260
gaagaaaagt ttttaaatgc tgaaagttac tataagaaag ctttggcttt ggatgagact 1320
tttaaagatg cagaggatgc tttgcagaaa cttcataaat atatgcagaa atctttggaa 1380
ttaagagaaa aacaagctga aaaggaagaa aagcagaaaa caaagaaaat agaaacaagt 1440
gcagaaaagt tgcgtaacgt cttaaaagaa gagaagaggc taaagaagaa aagaagaaaa 1500
tcaacttctt cttcaagtgt ttcttctgct gatgaatcag tgtcttcatc atcatcctct 1560
tcctcttctg gtcacaaaag gcataagaaa cataagagga accgttcaga gtcttctcgc 1620
agttccagaa ggcattcatc tagggcatcc tcaaatcaga tagatcagaa taggaaagat 1680
gagtgctacc cagttccagc taatacttca gcatcttttc ttaaccataa acaagaagtg 1740
gagaaactac tggggaagca ggataggtta cagtatgaaa agacacagat aaaagagaaa 1800
gatagatgcc ctctctcttc atcttcactt gaaataccgg atgatttggg agtgtactcc 1860
tatttattta aaaagttaac tataaaacag cctcaggcag gtccttcagg agatattcca 1920
gaagagggca ttgttatcat agatgacagc tccattcatg ttactgaccc tgaagacctt 1980
caagtgggac aagatatgga ggtggaagac agtggtattg atgatcctga ccacgggtag 2040
gcttaggttt atgtgtgtgt atgtgtctta gtttttaaca aaaaaattaa aaagtaaaaa 2100
aactaaaaat agaaaaatgc ttagagaata aggatataaa gaatattttt gtgcagttga 2160
acaatgagtg cttaagctaa atgtcatcac aaaagagtaa aaaaatttta caaaattaaa 2220
aatgtttaaa gttaaaaagc tctaggaagc taaggtcaat ttattattgg agaaataaaa 2280
ttatttttat gaatt 2295

<210> 91
<211> 1777
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2804990CB1

<400> 91

actaaatgtg ccagtacctg agattccagg aaaaactaaa aatttgtttt tgcaattagc 60
caaacatgta gcccagtctc tacaagtcgc ctatgtcatg tttgtggagg aaccgtaaca 120
ggagatcaat ggccatggga agcccgagaa ttggttccta tagacccggt tcctgatgaa 180
ttcccagccc aaaagaacca ccctgacaaa ttttgggttc tgaaagtctt attattggac 240
agtattgcat agctagagaa ggaaaaggat ttactcatcc tgtagggcag cttagttgtc 300
ttgggcaaaa gatgtataat ggtaccacaa aaacagttac atggtggagt tccaattaca 360
cagaaaaaaa tccattcagt aaatttccaa agttgcagac tgtttgggcc cacccagaat 420
tgcactggga ctggacggcc cccactgggt tatactggat atgtggacac agagcttatg 480
ctaagctgcc tgatcagtgg acaggtagct gtgtaattgg caccattaag ccatctttct 540
tcttagtgcc cataaaaaca ggtaaacttc tgggcttccc agtctgtgct tcctgtgaaa 600
aatgaagcat agccataggt gattggaaag acgatgaatg gccccctgaa aaatcttac 660
aatactatgg acctgccact tgggcacaag atggctcatg gggatatcga accccccatct 720
acatgctcaa ccgaatcata tggttacaag ctgttttaga aattattact aataaaacca 780
ctcaagcctt gactgttctt gcctggcaag agactctgat gagaaatgct atctatcaaa 840
atagactagc tcttgattac ttgctagcag ctgaaggagg agtttgtgaa aaatttgacc 900
ttactaatta ttgtctacac atagatgatc aggggcaagt agttgaggat atagttaaag 960
atataacaaa actggcacat gcacctgtgc aagtgtggca cggactcaat ctgggagcca 1020
tgtttggaaa ttggttccca gcaataggtg gatttaaaac tcttataata agagtaataa 1080
tagtaatagg aacctgctta ctgctccctt gtctgatacc tgtatttctc caaatgataa 1140
aaaacttcgt cgcttaacct tagttcacca aaatgcttca gcacaagcat actatataaa 1200
tcactatcaa tctattgcac gaaaatacat aagtagcaaa aataagagtg agaactccca 1260
ctgataaaaa gtgagagtct caaagagggg aaatgaggga agagagagac cctctcatat 1320
tgttttatat tgttttatac tcagtacctc ttttaagaaa aaaacaacaa ggaagtagaa 1380
ccaaagacaa gcagcccagc gccaggccca aaaccaggcc tgggcctgcc tggcctaaac 1440
ccagtagtta aaaatcaact cataacttag aaactgatgt tattcataga ttccagacat 1500
tgtatagaag aacattgtga aactccctgc cctgttctgt ttctctctga ccaccggtgc 1560
atgcagcccc tgtcacatac cgcctgcttg ctcaaatcaa tcatgaccct ttcatgtgaa 1620
atctttagta ttgtgagccc ttaaaaggga cggaaattgt gcattcgggg agctcggatt 1680
ttaaggcagt agcctgctga tgctcccagc tgaataaagc ccttccttct acaaaaaaaa 1740
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaag 1777

<210> 92
<211> 3181
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 168571CB1

<400> 92
gactccacca gccacccctg agctccaacc ttggggggat gcagagatcg gggaggccca 60
aggccacctc ctcaggccca aagggaggcc actgcaggca agatgaggtg gcccagagga 120
gccacatggg gcctggaggc catctgggcc gaggcagagc caccgtcaca cgtggatctc 180
agagacccac aggaagagga tggtggctgg gagggctgag gcagtccccc agagaactgt 240
cctgccaggc cacaaggtca gtgagaggga gcccggctcc caccccaggg accaggtgtc 300
ctttttcagg tccaggtggc agtgtccatc actccagggg cccccactgg ggaacttctg 360
agaagccagg gcatggagca agagggtgat acccacaaac tctttcctca tccttccctg 420
atttgaaatg gagggtggga tgcagaggag agaggtggaa agacaggtgc acagccccag 480
gcctctcccc caccccctgcc tcgcccacca tgtggctgga acggtgacca cccagcaaga 540
acccatctcc caccctcacc agtgatcccg gaaaaccagt cactggggtc ccgcggggca 600
tccgaaggcc caagggcaag tcacctttca gaatgtcccc tctgtgcagc ctttgccagg 660
aaggaagctg gaccgggcca gctgcatgga tgccaggacc tctaggccct gagcatcagg 720
gggtccagcc caggacaccc caagcctggg cccctctccc tgctgaaggg ctctggggag 780
cgcggggaga agccagcagg catggaggct gcccatcccc cagccatggc ctggggcccc 840
acgcagctct ctgcctcccc caggagaacc cccgcctcac agaagacttc gtgtcccacc 900
tggagaccga gctggagcag tcgaggctgc gggagacgga gacactgggg gcccttcggg 960
agatgcagga caaggttctc gacatggaaa agaggaacag ctcgctgccc gacgagaaca 1020
atgtggcgca gctgcaggag gagctgaagg cgctcaaggt gcgggaaggc caggcggtgg 1080
cctcgacgcg agagcttaaa ctgcagctgc aggagctctc ggacacctgg caggcccatc 1140
tggcccgcgg cggccgctgg aaggagtccc cacggaagct ggtcgtgggc gagctgcagg 1200
acgagctgat gagcgtgcgt ctgcgcgagg cccaggccct ggccgagggc cgcgagctgc 1260
ggcagcgcgt ggtggaactt gagacgcagg accacatcca ccgcaacctt ctgaaccgcg 1320

tggaggcgga gcgcgcggcg ctgcaggaga agctgcagta cctggctgca cagaacaagg 1380
ggctgcagac gcagctcagc gaaagccgcc gcaagcaggc cgaggccgag tgcaagagca 1440
aggaggaggt gatggctgtg cgactgcggg aggcggacag catggctgcg gtggccgaga 1500
tgcggcagcg cattgccgag ctggagatcc agagggagga aggccgcatc cagggccagc 1560
tgaaccactc ggactcatcg cagtacatcc gcgagctcaa ggaccagatc gaggagctga 1620
aggccgaggt gcggctgctg aagggcccgc cgcccttcga ggacccgctg gctttcgatg 1680
ggctgagcct ggcgcggcac ttggacgagg actcgctgcc gtcgtcggac gaggagctac 1740
ttggcgtagg cgtgggcgct gccctgcagg acgcattgta ccctctgtcc ccgcgcgatg 1800
cgcgcttctt ccgccgtctg gagcggccgg ccaaggacag cgagggcagc tcagacagcg 1860
acgccgatga gctggccgcg ccctacagcc agggtctgga caactgaggc catgcccagc 1920
gcgcccggag tcaggaggcc gcagccgcgg ggggcgcccg ggcagtccgc gttctgctcc 1980
ccacctgccg cacttgacaa actacgcgcc ctctgtggct cggccacccc taaagcgagg 2040
cccggcgagg cagcgcagag ggtagggtcc gacctgggct cctcagggcc ccggggcagg 2100
ctctctctcc ccagcagtgt ttacccatct tggtctgtac cctccgggc cctctggcgt 2160
tccaggggtg cctggagggg ctgactgctc tcttaacagg agggcagagg gcaggggaca 2220
gacgacccag aggtcccagc actgaatgag caggcagctc ccacctcctg gcaggcttcc 2280
ttctgggaca ggggcgacat tgctgggaag tgctcaggag gtagccgagg cctgaggaag 2340
gagagcgcca gctctgggct ggacatcagc accccacaac actcctcggg atgaagtgac 2400
ccttgactag cccctggcca tctctagggg agtcaggccg ctggggacag atggccaggc 2460
cggcctctcc tgcctggcgc aggcaccgtg gcccctgcag cggaaaccaa agtcccccta 2520
ctatgtgcgg ggcgcctggg gactgagtgg ctaacgggaa gcctccctgc ttctctgggg 2580
ccagagcagc ttccaggaag tgaagagacc cctctcccca gtgccccaca tctttctctg 2640
ggagacactc gcgcccccta tcctggccag tgatggaggg gtatctctac aggggtcctc 2700
ggtctccatc tacttccctt tcttcatgtg tttccagccc caccctcagc cagagccagg 2760
cccccaggca ggagcttccc agcgagcggc ctcccctcac ttcctcctgg tggtcccgtt 2820
gtctctgctg aatcagagct gagaacggtg ccaaaatcca acgatgcccc ccaggcccct 2880
tcctcccttc ccccggcccc ccgggcctca tgaccctggg ggctgccccc ctcggtgctc 2940
ccggagcctc cagtagagta gcgtcacaag caatctccct ggcgcttcct gggtggggac 3000
ccctgctccg tccccgcttc ctagctgccc acttttcagt gttacgaagc ctggggaccg 3060
gggcaggcac ccacggggct ctccacacgc cccctacact gcccgccacc attttgcaca 3120
ctgcctgttc acatgtcgcc caggcgggaa aaatggaaaa taaagtgtat ttacacagtc 3180
a 3181

<210> 93
<211> 5987
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1286391CB1

<400> 93
ccagtgctgc caacagggca agggaagccc gtggctcaag gggggctgac agatcagtca 60
gtacctggag ctcaaacagt gcctgaaact ctcaaagtgc ccatggctgc agcagtgccc 120
aaagctgaaa atccctccag aactcaagtg ccatctgcag ctcccaaact gcctacatct 180
cgaatgatgc tggcagtgca cacagagcct gcagctcccg aagtgccttt ggctccaaca 240
aagccaacag ctcaactgat ggccacagct caaaaaacag ttgtgaatca accagtgttg 300
gtagctcaag tggaacccac aactccaaaa actccccagg ctcagaagat gcctgtagca 360
aaaacatcac ctgcaggtcc caaaacaccc aaagctcaag ccgggcctgc agctacagtt 420
tccaaagcac ctgcagcttc caaagcacct gcagctccca aagtacctgt gacccccaga 480
gtctccagag ctcccaaaac acctgcagct cagaaggtgc ccacggatgc agggccaacc 540
ttggatgtag ccagacttct gagtgaggtc cagcctacat caagggctag tgtctcctta 600
ctgaagggcc aggggcaggc tggaaggcag ggtccccagt ccagtggcac cttggccctc 660
agcagtaagc accagtttca gatggagggg ctcctggggg cttgggaggg ggccccaagg 720
cagccacctc gccacctgca agcgaacagc acagtgacca gcttccagag gtaccacgag 780
gccctgaata caccccttcga gctgaacctg tcaggggaac ctggaaacca ggggttgcgg 840
cgagtggtca tcgatggcag cagtgtggcc atggtgcatg gcctgcagca cttcttctcc 900
tgccgaggaa ttgccatggc agtgcagttt ttctggaacc ggggacaccg agaggtcact 960
gtgtttgtac cccacctggca gctgaagaag aaccggaggg tgagagagag ccactttctg 1020
acgaagctac actcgctcaa gatgctttca atcacaccct cccagcttga gaatggcaag 1080
aagatcacca cctacgatta taggttcatg gtaaagctgg cagaggagac agatggcatc 1140
attgtcacca atgagcagat tcacatcctg atgaatagtt ccaagaaact gatggtcaaa 1200

gatcgcttgc tgcctttcac ctttgcgggg aatctcttca tggtgcctga tgacccctg 1260
ggccgtgatg gccccacctt ggatgagttt ctgaagaagc caaacaggtt ggacactgac 1320
attggcaact tcctgaaggt gtggaagacc cttcctccca gctcagccag tgtcactgag 1380
ctgagtgatg acgctgactc tgggcccctg gagagtctgc cgaatatgga agaagtcagg 1440
gaagagaagg aggagaggca ggatgaggag cagagacagg ggcagggcac acagaaggcg 1500
gctgaggagg acgaccttga ctcttcgctg gcgtcagtgt tcagggtgga gtgcccgtcc 1560
ctttcggagg agatcctgcg gtgcctcagc ctccatgatc cccctgatgg ggccctggac 1620
atcgacctcc tgccaggggc agcttctccc tacctgggca tcccctggga tggaaaggct 1680
ccctgccagc aggttcttgc ccacctggcc cagctcacca tccccagcaa cttcaccgca 1740
ctctccttct tcatgggctt catggactcc cacagggatg ccatccctga ctatgaagcc 1800
ctagtgggcc ccctgcacag cctcctcaag cagaagcctg actggcagtg ggaccaggag 1860
catgaggagg ccttcctggc cctgaagcga gccctggtgt ctgccctctg cctgatggcc 1920
cccaactccc agctgccctt ccgcctggag gtgaccgtga gccacgtggc cctgacggcc 1980
atcctccatc aggagcactc agggaggaag caccccatag cctatacctc aaaacccctc 2040
ctccctgatg aggagagcca gggcccccag tcagggggtg acagccccta tgctgtggcc 2100
tgggccctca agcatttttc ccgctgcatt ggagacaccc cggtggtcct ggacctttcc 2160
tatgcctccc ggaccactgc ggaccctgag gtgcgggagg gccgcagggt ttccaaagct 2220
tggttgatcc gatggtccct cttggttcag gacaaaggca agagggccct ggaattggcc 2280
ctcctccagg gcctgctggg ggagaaccgc ctgctcaccc ccgcggcctc catgcctcgc 2340
ttcttccagg ttctgccgcc tttctctgac ctgtccacgt tcgtctgcat ccacatgtcg 2400
ggctactgct tctaccgtga ggatgagtgg tgtgctggct ttggtctcta tgttctatcg 2460
cccaccagcc cccctgtctc cctttccttc tcctgctccc cttacacgcc aacctatgcc 2520
cacctggcag ccgtggcctg cggcctggag cgctttggcc agtccccact cccagtggtt 2580
ttcctcactc actgcaactg gatcttcagc ctcctgtggg agctcctgcc cctctggagg 2640
gctcggggct tcctctcctc tgatggggct ccactccctc acccaagcct gctctcctac 2700
attatatccc tcacctctgg cctctcatcc cttccgttta tctaccgaac ctcctaccgg 2760
ggctctctgt ttgctgtgac agtggacacc ctggccaagc agggtgccca gggggtgggg 2820
cagtggtgga gtttgccaaa ggatgtgcca gcccctacag tgagtcccca tgccatgggc 2880
aagaggccca atttgctggc attacagctg agtgacagca ccctggccga catcattgcc 2940
aggctgcagg ctgggcagaa actgtctggc tcctcaccgt ttagttctgc ctttaactca 3000
ctcagcctcg acaaggagag tggcctgctt atgttcaagg gagataagaa gcccagggtc 3060
tgggtagtcc cgacgcaact ccggagggat ctgattttct ctgtgcatga cattcccttg 3120
ggggcccacc agaggcccga agagacctac aagaagttgc gtttgctggg gtggtggcct 3180
gggatgcagg agcatgtgaa agattactgc aggagctgct tgttctgcat cccccgaaat 3240
ctcataggca gcgagttgaa ggttattgag tccccatggc cctcaggtc gaccgccccc 3300
tggtcgaacc tgcagatcga ggtggtgggc ccggtcacca taagtgagga gggccataag 3360
catgtactta ttgtggctga cccaaacacc aggtgggtgg aggcattccc cctgaagccc 3420
tacacacaca cggctgtggc ccaggtgctg cttcagcatg tgtttgcaag gtggggtgtt 3480
cctgtgaggc tggaggcagc ccaggggccc cagtttgccc ggcacgtcct tgtgagctgt 3540
gggctggccc tgggagccca ggtggcctcc ctgagtcggg acctccagtt cccctgcctg 3600
acgagctcag ggcctactg ggaattcaag agggccctca aggagttcat cttcctgcat 3660
gggaagaagt gggcggcctc cctgcctttg ctgcacctgg ccttcagggc ctcctccact 3720
gatgccacac cgttcaaggt cctgaccggg ggtgagtcaa ggctcacgga gcccctgtgg 3780
tgggagatga gcagcgcaaa cattgaaggg ctcaagatgg acgtcttcct gctacagctg 3840
gtgggggagc tgctggagct ccactggagg gtggctgaca aggcgagtga aaaggccgag 3900
aacaggcgtt tcaagcggga gagccaggag aaggagtgga atgtgggtga ccaggtcctt 3960
ttgctgtccc tccccaggaa tggcagcagt gccaaatggg tgggtccctt ctatatcggg 4020
gaccggctga gcctgtcact ctataggata tggggcttcc caaccccaga gaagctgggg 4080
tgcatctatc ccagcagtct gatgaaggcc tttgccaaga gtggcacccc gctgtccttc 4140
aaggtcttgg agcagtgagc gggagcagcg gggtgcccc ctgccccagg gccgtgggtt 4200
tctgctgcta ggcctccccc tgtcccagca gtgctctcag tccactgggg gccctcagtt 4260
gtgccttttg tagagaactt gcttcataaa gctttgctga attgccttga actaggacc 4320
agcatcccca tggaaacatc cccagtttgg ggtacttgga gaatttgcca aaggctgtca 4380
gatatgggtc cctggcagtt ttacactggg aaatagagtg ctcctctagg ctataccagg 4440
ctcagtgtct tctccccaag tatcctcttt ccccttcaca tgtaggtgtg tggtggtgtg 4500
cacacacact cacgaaagaa tctatcttgg ccatgaaact gtggctgttg accttggaat 4560
tggaaccaca gtcctttccc atacagaaac cccaaatgtg ggctcctccc tccaacctgt 4620
tcttttgggg gaccctttgc ccttagagct gtcacagatg acataagcct gggcaggctg 4680
tggggaggcc gctgcctcct agcctcacgg tcagctttct caagccaggt gtggcctgca 4740
cagctgtcag ggcagggctg gccatcctcc caggcctcaa gggcagtgtc ttggagtggg 4800
aggatggcca gccacaagcc accagcttgt cagcatggga agggcaaggg ggaaatgggt 4860
tggcctacct catttgactc cagccaatgg agacaattcc tgacccctgc tttgatggtg 4920
ttgaccccac aggaatcaag gatcttgatg ccaagtgtgg catctctacc tgaagagctg 4980

cttctgttag acccaggggc cccggcctct gttttaaggg ggcagggcgt ctgcaacagg 5040
agtggcacac ggtgaagtgc tggcatggct ctacctccca acccctccca accccatccc 5100
aaagcctaca gtcctctgct ccttctacct ccactgtatc ctgcatccca gaccctacag 5160
attgtgtgat tgcttcatct gtatcacccc ccgagtcctg tggacctgcc ttctgtgtag 5220
agcaaaaccc aggctccctc acagccattc tttgcagaga tcacccttgc agtgagtgtg 5280
agttccttcc aggtgtgtgc gcgtacactc accctctgaa ttgctgccgc tgccaaggaa 5340
gtgggctcca ggtgaaggcc agtgccacgt cactctggct ggccttctta gtagtttcat 5400
ttctccggaa gctgagccag tctcctggtc tagcccaggt tgccagaacg cttggcattg 5460
cagagtgcta gagccagtgg agaacttgcc aacttgattg ttttacagca gaggaaagag 5520
gatcacagag ggaaagtgat tcacccaaag tcacacagca agttcatggc tgagctgaga 5580
ccaggattaa gcttcctgac tcccagttcg ccatgaaaag ggttctggca acaggttcaa 5640
gctggagaat ccttcaaaat gctacaccca cattctctcc aactcttcat ctccctgatc 5700
ttccagacaa actacctgga tgttgccctt aaaccatttc tagctgttaa ccctgtccag 5760
aaaaatgatt gagtgatagc tgagaagtgg aaagtgtggg atttttggca ggtgctctct 5820
ttcctccgcc ccccgcgcca ttctttctct tcctcctctc tgtaatggta tgtccagcct 5880
cactctccct ccctggtgct gtatgcgttc cccctgttag ctacatttgt gatcacatac 5940
ccttctttta agtgaatttt tttcatttga tttgtcaata aacgaat 5987

<210> 94
<211> 765
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2007684CB1

<400> 94
gtcggggctc tgaggctccg gcctgacctc tcctcggggt cgacgggaag gtctccggat 60
gccaggactc gcaaagggcc gaccagaatg aggaaatccc aggcggagtc cgggggaggc 120
agcacggcat cccagcctca gggctgcccg gacggtgttg gttggggtaa gattggggaa 180
gatgccaggt gatagaatcc accctgcctc tgacagcatc caatctggag caaaggcttc 240
tttcaacctt cccccaagct cccaatacaa gcccaaattc tggaacaagt cctttccaac 300
acctccttca gagaagctcc ctggctccag gcagtgcttg ctcccctcac tgcagcagga 360
agcctaactt gcgaactgca ggtgtgctcc aggtggtcac tggatgcaag aaatggtgcg 420
agaactttgg atgtggaacg tggaggagga ggaacacgaa gtggggatct gcacttgggg 480
tggtcagcac tgcggatgcc cagcaaagtc actgcctggt ccacatcccg gaggggtctc 540
tgcgcctcag tcggcatcgc agctgatggt gaaacttttg gtgtggcaga agagtgtcca 600
caaacttcgg aagttgctag agaaaactga aaactactag ttttgaaagc ctcacttctg 660
ccaatggaag cacattccag catgtcgcca acgcaatcca ctttccacca ctttcacaaa 720
aaacgttact gcacaattat actatcctac ccttaaaaaa aaaaa 765

<210> 95
<211> 2674
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2227040CB1

<400> 95
gcgacctcct aaatcaaagt tgggcggcga ggactgggcg gaggggccgc cgggtgccgg 60
gctggggccg gcgggggagg ggaggggatg gggggcgggc gcagcggacg gcggagccgg 120
ctggacacgg acagctcctt ggctccctcc ctcggctcat ccccgcggcc ccacttccct 180
cctccccgcg ctgcgagcag catcctctgc agaccctcgg tcctcgcgcc cggggtcgtc 240
ccgctcctgc ggctcagcgt ggtggcctcc cctcgcccgc caccccggca actttctctg 300
cccgctcccg cgggttgggg gctgccgtgc cgggggctaa ctgggggcag cctctggaga 360
acggcttcag agtcccggag acgcccgcca cccgcagcct gcccgcggtg ggcctgccgt 420
cggatctcgg caccctctcc tcccctctcc ccgaaccatg accgagatga gcgagaagga 480
gaacgaaccg gatgacgcgg ccacccacag ccccccaggg accgtctccg ccctccagga 540
aaccaagctc cagcgattca agcgctccct ctccctcaag accatcctcc gaagtaagag 600
cttggagaac ttcttccttc gctcgggctc tgagctcaag tgccccaccg aggtgctgct 660

gacgcccccca accccactgc cccctccctc cccaccaccc acagcctcgg acaggggcct 720
ggctacccca tccccctccc catgcccagt cccacgcccc ctggcagcgc tcaaaccagt 780
gacgctgcac agcttccagg aacatgtctt caagcgagct agcccttgtg agctgtgcca 840
ccagctcatc gtaggaaact ccaaacaggg cttgcgatgt aagatgtgca aagtcagcgt 900
ccacctctgg tgctctgagg agatctccca ccagcaatgc ccaggcaaga cgtccacctc 960
cttccgccgc aacttcagtt cccctctcct ggtgcatgag ccgccaccag tctgtgccac 1020
aagcaaagag tccccaccca ctggggacag tgggaaggtg gaccctgtct acgagaccct 1080
gcgctatggc acctccctgg cactgatgaa ccgctccagt ttcagcagca cctctgagtc 1140
cccgacaagg agcctgagtg agcgggatga gctgaccgag gatggggaag gcagcatccg 1200
cagctctgag gaggggcctg gtgacagtgc atctccagta ttcacagccc cagcagagag 1260
tgaagggcca ggaccagagg agaagagtcc tggacagcag ctccccaaag ccaccctgcg 1320
gaaggatgtg gggcccatgt actcctacgt tgcactctac aagtttctgc cccaggagaa 1380
caatgatctg gctctgcagc ctggagatcg gatcatgctg gtggatgact ctaacgagga 1440
ctggtggaag ggcaagatcg gcgaccgggt tggcttcttc ccagctaatt ttgtgcaacg 1500
ggtgaggcca ggcgagaatg tttggcgctg ctgccaaccc ttctccggga acaaggaaca 1560
gggttacatg agcctcaagg agaaccagat ctgcgtgggc gtgggcagaa gcaaggatgc 1620
tgacggcttc atccgcgtca gcagtggcaa gaagcggggc ctggtgccag tcgacgccct 1680
gactgagatc tgagaggagc caagggaacc cagatgacac ccttgcccat gcctggacct 1740
tgcttctggc cagggagggg atcaggcccc cttgtccact ccataccctt cctccctctg 1800
tccctctcct aggtgccact taccgtggct taggagcctt ttgtactggg gaagatttta 1860
tttcttgggt ggggtgccct gagtgggttg atcctctggg acttggcggg gatggggtgg 1920
ggtgggaggg aatgaggaag ctacaggtag gtccacccaa cgcccaggca ccccagtcta 1980
cttgctgggc tgagtccagc cctggggaag attcctgggg aatttatttg ctgcttctcc 2040
cagccttccc ctgcccacac cacctgcgca tgcctggcac aaccctctcc ccacagagcc 2100
tctgcttggg tctatgtgtg tgtgaccgtg catctgctcc tttgcaagtg gggtcttggg 2160
agcaggcctt tttgtgtcct cggccctcca cccttggtgg ggaggggacc caggaccata 2220
ggaaggtcct gagtccacct cctggtctcc acctctgtca ttccgtcaat tctcaggaaa 2280
gttttgcagg aatctctccc gttacttgaa acctggcgg acagatgagg ggaaggctga 2340
ggtccctgga aggtgtaaca aatcagagac tatctgttca agtagaaaac ccaagagctg 2400
gccaggtgtg gtggctcatg cctgtaatcc cagcactttg ggaggctgag gtgggcagat 2460
cacctgaggt caggagttca agaccagcct ggccaacatg gtgaaacccc ctctctacta 2520
aaagtacgaa acttagccag gcatgatggg ggtctgcctg taatgccagc tactccagag 2580
gctgaggcaa gagaatcact tgaacccggg aggcggaggt ttcagtgagc cgagattgca 2640
ccattgcacc ccagcctggg caacagagcc gaga 2674

<210> 96
<211> 5920
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4346130CB1

<400> 96
cgccaggcag cagaggcgac tgcggacagc cggcggggac cgggaaagag gaggcctggc 60
gcctccggca gaggggcgag gaccgagcgc ccaaggctgg gagagacgcc ggcgctgagc 120
ttcccagcgc cgcctcacct cgcttttcga actccgcgac ccgagcgcaa ggcttaagca 180
ggtaaagcag tcactgtgaa gaaaataaca ttttaagaaa cttttctgtc acaattgata 240
gaggaaaagc cagtagagaa gtacatcttc ttgtgtgtgt gtctcatcgt gccacttttt 300
gtaactcagc ttagtaatga agccaccatt atgtcctcag tcagtgaagt aaatgtggat 360
ataaaagatt tcctaatgag cattaatttg gagcagtatc tcttacattt ccatgagtct 420
ggttttacta ctgtgaagga ctgtgcagca ataaatgaca gcctgctgca gaaaattgga 480
atatcaccta caggtcaccg taggaggata cttaaacagt tacagataat cttgtcaaaa 540
atgcaagata ttccaatata tgcaaatgtt cataaaacta agaagaatga tgacccttca 600
aaggattacc atgttccatc ttctgatcag aatatctgca tagaactttc caattctggt 660
agtgttcaga catctagccc accgcagttg gagactgtta gaaaaaatct tgaagacagt 720
gatgcaagtg tggaaagaag ccagtatcct caatcagatg ataagctgtc tcctcctaaa 780
cgcgacttcc ccactgcaga ggaaccacac ctgaatttgg gttctttgaa tgattcttta 840
tttggtagtg acaatattaa aatagaatca ttgattacaa agaagactgt ggatcacaca 900
gttgaagaac aacaaacaga aaaagttaaa ttgatcacag aaaatctcag taagctccct 960
aatgcagact ctgaatgcct ttcttttgtt ggctgttcaa catcaggaac aaattctgga 1020
aatggaacaa atggtttatt agaaggatca ccaccatccc cattctttaa gtttcaagga 1080

gaaatgattg taaatgactt gtatgttcca tcatcaccaa tcctagcacc tgtgagaagt 1140
cgtagcaagt tggtttcaag accatctcga tcttttctgc taagacatcg acctgtacca 1200
gagattccag ggtcaacaaa aggagtttct gggagctatt tccgtgaaag aagaaatgtt 1260
gctacctcaa ctgaaaaatc tgtggcatgg caaaattcaa atgaggagaa ttcatcttcc 1320
atctttcctt atggagagac ctttctcttc cagagactag aaaattccaa gaagcgatct 1380
ataaagaatg aatttttgac ccagggagaa gcactcaaag gggaagcagc tactgcaaca 1440
aactctttta tcatcaaatc aagcatatac gataacagaa aggagaaaat aagcgaggac 1500
aaggtggaag atatttggat acctcgagag gacaaaaaca atttttgat agacactgct 1560
tctgaatcag aatactcaac agtagaagaa tgctttcaga gtttaagaag aaaaaattca 1620
aaggcatcta aatctaggac tcaaaaagcc ttgattttgg actccgttaa taggcacagt 1680
tatccgttaa gctcaacaag tggaaatgct gattcatcag ccgtttcttc acaggcaata 1740
tctccctatg cctgctttta tggagcatct gcaaagaagg ttaaatcagg atggctggat 1800
aaactctctc ctcaaggaaa acgcatgttt caaaagagat gggtgaaatt tgatggcctt 1860
agcatttctt actacaataa tgagaaggag atgtattcga aggaataat tcccctttct 1920
gctatatcaa cagtacgagt tcaaggagac aacaaatttg aagttgttac aacacaaaga 1980
acttttgttt ttagagtaga aaaagaagag gagagaaatg actggatcag catactatta 2040
aatgcactga aatcacaatc ccttacctcg cagtctcaag ctgttgttac acctgagaaa 2100
tgtggatatc ttgaattgag aggctataag gcaaaaattt ttactgtgtt aagtggaaac 2160
agtgtgtggc tttgcaaaaa cgaacaggat tttaagagtg gacttggtat taccataatt 2220
cctatgaatg tagcaaatgt aaagcaagtg gaccgaactg tgaaacaatc ttttgaaata 2280
atcactccct acaggagttt cagctttaca gccgagactg aaaaggagaa acaagactgg 2340
attgaagctg tgcagcaatc aatagcagaa actctctctg attatgaagt agctgagaag 2400
atttggttca atgaatccaa caggagctgt gcagattgta aagccccaga tcctgactgg 2460
gcatccatca atctctgtgt tgtcatctgt aagaagtgtg caggacagca tagatcttta 2520
ggaccaaaag attccaaggt tagaagtcta aaaatggatg ctagcatttg gagcaatgaa 2580
ctcatcgagc tttttattgt cattggaaac aaaagagcaa atgacttttg ggctggtaat 2640
cttcaaaagg atgaagaatt acatatggac tcaccagtag aaaagagaaa aaactttatt 2700
actcagaaat ataaagaagg aaaattcaga aaaactcttt tggcatctct caccaaagaa 2760
gaattaaata aggctctatg tgctgctgta gtgaaaccgg atgttctaga aacaatggct 2820
ttgctgttca gtggagcaga tgtcatgtgt gccaccggag accccgtgca tagcaccccc 2880
tatctgctag ccaagaaagc tgggcaaagt ctgcaaatgg aatttctcta ccataacaaa 2940
ttctcagatt tccctcaaca tgatattcat tccgagggtg tattaagtca agagtcttcc 3000
cagtccacat tcctctgtga ctttttatat caagctcctt ctgctgcttc taaactctct 3060
tcagagaaaa aactgcttga agagacaaat aaaaaatggt gtgttttgga aggaggcttc 3120
ttgagttact atgaaaatga taagtctacc acacctaatg gcaccattaa tatcaatgaa 3180
gttatctgcc tggctataca caaagaggac ttctatttaa atactgggcc catctttatc 3240
tttgagatct acttaccctc cgaacgtgtg tttttatttg gagctgaaac atctcaagct 3300
caaagaaaat ggacagaggc aatagccaag cattttgttc ccttatttgc tgaaaactta 3360
acagaagctg actatgattt gattggtcaa ctcttctaca aagactgcca tgccctggat 3420
cagtggagaa aaggctggtt tgctatggac aaatccagct tgcatttttg ccttcaaatg 3480
caagaagttc agggagatag aatgcactta agaagactgc aagagctaac aatcagcaca 3540
atggttcaaa atggggaaaa actggatgtt ttactcttgg tagaaaaagg gagaacatta 3600
tacatccatg ggcataccaa gttggatttc acagtctggc atactgcaat tgaaaaagca 3660
gcaggtacag atggtaatgc tttacaagat cagcagctca gcaaaaatga cgttcccatt 3720
atagtgaaca gctgtatagc atttgttaca cagtatggtt taggatgcaa atatatctat 3780
caaaagaatg gtgatccttt gcatataagt gaactcctgg agagtttcaa aaaggatgca 3840
agaagcttta aattgagggc tggaaaacat cagcttgaag atgtgacggc tgtgttgaaa 3900
agttttctct ctgacattga tgatgcactg cttactaagg agctctaccc atattggatc 3960
tctgctttag atacgcaaga tgacaaggaa agaattaaaa aatatggagc atttatacgt 4020
tctcttccag ggtcaaccg agcaacacta gcagctatca ttgaacacct gtatagggtt 4080
cagaaatgct cagaaatcaa tcacatgaat gcccataatt tggccttggt cttttcatcc 4140
tgtttgtttc aaacgaaggg acaaactagt gaagaagtga atgtaattga ggacctaatt 4200
aataattatg tagaaatatt tgaggttaaa gaagatcaag tcaaacaaat ggacatagaa 4260
aatagcttta ttaccaagtg gaaagacacc caagtttccc aggctggaga tttgttaatt 4320
gaagtatatg tagaaaggaa ggaacccgac tgtagtatta taattcggat atctcctgtg 4380
atggaagcag aagaattaac taatgatata ttagcgataa aaatattat tcctacaaaa 4440
ggtgatattt gggccacatt tgaagtcatt gaaaatgaag agctagagcg tcctcttcac 4500
tacaaggaaa atgtactgga gcaggtgctt cggtggagtt cattagctga acctggctct 4560
gcttacctgg tggtgaagag attcttaacc gctgacacaa ttaaacactg cagtgaccgg 4620
agtacactgg gaagcatcaa agaaggaatc ttgaaaatca aagaagaacc atccaaaata 4680
ctatctggaa ataagtttca agaccggtat tttgttttac gagatgggtt tctctttctt 4740
tacaaggatg tgaagagtag taaacatgac aagatgtttt ctctcagttc catgaagttt 4800
tatcgtggag tgaaaaagaa aatgaagcct ccaacaagct ggggattgac cgcatattct 4860

gagaaacatc actggcacct gtgttgtgat agttcacaaa ctcagacgga gtggatgacc 4920
agtatcttta ttgcccagca tgaatatgat atatggccac cagctggaaa ggaacgaaaa 4980
cgttcaataa ccaaaaatcc caaaattgga ggtttgcctc tgattcctat acagcatgag 5040
gggaatgcaa ccttggcccg gaaaaatatt gagagtgcaa gagcagaact tgaaaggctg 5100
cggctcagtg aaaagtgtga taaagagtcc gtggactcta gcttaaagga gagagcttcc 5160
atggtggccc actgcctgga gcacaaggac gataaacttc gaaatcgacc ccgaaaacat 5220
cggagtttca actgcctgga ggacacagag cctgaagccc cacttgggca accaaaaggc 5280
cataaaggcc taaagacatt gaggaagact gaggacagga atagcaaagc tactttggac 5340
tctgatcata agttaccatc aagggtaatt gaagaactta atgtggttct acaacggtca 5400
agaacccttc caaaagaatt acaggatgag cagattttga agtaggaaat aaaatgaatt 5460
gctccccaga ttattttttt aatattgtat gcaatcttta tgtgtaactc caaaactgaa 5520
ctataagata attcataaga atttgcctat tgataggcaa atttgtctat gcttaggcat 5580
ttaaggagca tttaaaaata tgtatatatg tgactgaatg taagattaac tttattttgg 5640
tctgaacaaa ccttgtaaag ttccattgta ttaacgtgtg taagaagaca ttgggttact 5700
taggtggtca gtccttttaa agtagttgtc tgttaagtgg tataatttat gagcagaaat 5760
ctcaaactgt actgtattgt ataaaatgct gtgtttaaat gaagcctaag aaactatctg 5820
taacatattt tgcactttga gaaaccttgt atattaaatt accatatgtt tttaggttta 5880
cagaagttct accttaggga aaagaaaggg aagtaaattg 5920

<210> 97
<211> 1689
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 55117040CB1

<400> 97
ctggtttctc ctggctgtta aaagcagatg gtggctgagg ttgattcaat gccggctgcc 60
tcttctgtga agaagccatt tggtctcagg agcaagatgg gcaagtggtg ctgccactgc 120
tttccctgct gcagggggag cggcaagagc aacgtgggca cttctggaga ccacaacgac 180
tcctctgtga agacgcttgg gagcaagagg tgcaagtggt gctgccactg cttcccctgc 240
tgcaggggga gcggcaagag caacgtggtc gcttggggag actacgatga cagcgccttc 300
atggatccca ggtaccacgt ccatggagaa gatctggaca agctccacag agctgcctgg 360
tggggtaaag tccccagaaa ggatctcatc gtcatgctca gggacacgga tgtgaacaag 420
agggacaagc aaaagaggac tgctctacat ctggcctctg ccaatgggaa ttcagaagta 480
gtaaaactcg tgctggacag acgatgtcaa cttaatgtcc ttgacaacaa aaagaggaca 540
gctctgacaa aggccgtaca atgccaggaa gatgaatgtg cgttaatgtt gctggaacat 600
ggcactgatc caaatattcc agatgagtat ggaaatacca ctctacacta tgctgtctac 660
aatgaagata aattaatggc caagcactgc tcttatacgg tgctgatatc gaatcaaaaa 720
acagcatggc ctcacaccac tgctacttgg tatacatgag caaaacagca agtggtgaaa 780
ttttaatcaa gaaaaaagcg aatttaaatg cgctggatag atatggaaga actgctctca 840
tacttgctgt atgttgtgga tcagcaagta tagtcagccc tctacttgag caaaatgttg 900
atgtatcttc tcaagatctg gaaagacggc cagagagtat gctgtttcta gtcatcatca 960
tgtaatttgc cagttacttt ctgactacaa agaaaaacag atgttaaaaa tctcttctga 1020
aaacagcaat ccagaacaag acttaaagct gacatcagag gaagagtcac aaaggcttaa 1080
aggaagtgaa aacagccagc cagaggcatg gaaactttta aatttaaact tttggtttaa 1140
tgtttttttt tttggcctta ataatattag atagtccaaa tgaaattacc tatgagacta 1200
ggcttggaga atcaatagat tctttttttta agaatctttg ggctaggagc ggtgtctcac 1260
gcctgtaatt ccagcacctt gagaggctga ggtgggcaga tcacgagatc aggagatcga 1320
gaccatcctg gctaacacgg tgaaacccca tctctactaa aaatacaaaa acttagctgg 1380
gtgtggtggc gggtgcctgt agtcccagct actcaggagg ctgaggcagg agaatggcat 1440
gaacccggga ggtggaggtt tgcagtgagc cgagatccgc aatacactcc agccctgggt 1500
gacagagcaa gactcttgtc tcacaaaaaa cataaaacac aaggggggc cgtcaacatg 1560
cttttaaagg gcccaatttc gccttatagg gagtcggttt tacacaaatt caggggcccg 1620
gcgttttaac aacgtccgga attgggaaac cccttgggtt taccccaatt taaatcgcct 1680
tggaagatc 1689

<210> 98
<211> 1314
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7472392CB1

<400> 98
gcgagctgcg cgtggcctgg gctccttccg aatggtggcg cctctgagcg ccagttcgag 60
gaggagtcgg aggaggagcc tgaatgtttg gagatagact tcaagtcccg gaccttatcc 120
gtgcgccgct tcggtttgca gggactgtac cagctcgcca tggatatcat cataatgatc 180
cgagtgtgta aaatgttccg ccaaggcctc aggggattcc gggaatatca aatcattgag 240
actgctcact ggaagcaccc tatcttctcc ttctgggata aaaagatgca aagccgagtc 300
acatttgata ccatggactt cattgcagag gagggtcact ttcctccaaa ggccattcag 360
atcatgcaga agaagccttc ctggagaaca gaggatgaga tccaggccgt ctgtaacatc 420
ttgcaggttc tggatagcta tcggaactac gcagagcccc tgcagctgct cctggccaaa 480
gtcatgcgct ttgaacggaa atggacgttc tgcatgcttc agtgaggagg tccaccatcg 540
tctgtatgga agaaacggag ttcctggttg ttgaccggga ggacttcttt gctaataagc 600
tggaccagga agttcagaag gatgctcagt atcggtttga attttttagg aagatggagc 660
tgtttgcatc atggtctgat gagaagctct ggcagctggt agccatggcg aagatagaga 720
ggttctcgta tgggcagctg atctcaaaag attttggaga gtcacccttc atcatgttta 780
tcagcaaggg cagctgtgaa gtcctgcggc tgttggacct tggggcctcc ccttcctacc 840
gtagatggat ctggcagcac ctggagctga tagatggcag acctctgaag acccacctga 900
gtgaatactc tcctatggaa agatttaagg aattccagat caaatcatat cctctgcaag 960
actttagctc cttgaaactt ccacatctca aaaaagcctg gggctacag gggacaagct 1020
tcagcaggaa gatcagaacc tcaggagaca ctctccccaa gatgctgggc ccgaagatcc 1080
aatccaggcc tgctcagtcg atcaaatgtg ccatgatcaa tatcaagcct ggtgagctcc 1140
ccaaggaggc tgcagtgggg gcctacgtga aggtgcacac tgtggagcag ggagaaattt 1200
tggtgagtgt gccaagagct ctgttcacca tggagtacgt aacatgaatg ggctggaatg 1260
ttaagtaaat gacagtaaag gttttacaaa caaaaaaaaa aaaaaggggc ggcg 1314

<210> 99
<211> 3322
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4028960CB1

<400> 99
ggagccccca cccaggcgca tggctccatg aagcaagagg aggaggcaga gcgcgagagg 60
tcccgtccgc cccagccgcg tctgggccag gagaaaaggt acaggaagaa ggaatggaat 120
cggaaggaag aaaattcatt ccagagacaa aatagctccc ctcctctgtt gtagctgttc 180
ctctggcaat acaaactttc tgctttcagg ttctgcagca ttccagacct tgagcactga 240
atcaggatgg gaaaaagacg ttgtgttcct ccactcgagc ccaagttggc agcaggctgt 300
tgtggggtca agaagcccaa attatctgga agtggaacgc acagtcacgg gaatcagtcc 360
acaactgtcc ccggctctag ttcaggacct cttcaaaacc accagcatgt ggacagcagc 420
agtggacggg agaatgtgtc agacttaact ctgggacctg gaaactctcc catcacacga 480
atgaatcccg catcgggagc gctgagccct cttccccggc ctaatggaac tgccaacacc 540
actaagaatc tggtggtgac tgcagagatg tgctgctact gcttcgacgt actctactgt 600
cacctctatg gcttcccaca gccacgactt cctagattca ccaatgaccc ctatccgctc 660
tttgtgacgt ggaagacagg gcgggacaag cggcttcgtg gctgcattgg gaccttctca 720
gccatgaatc ttcattcagg actcagggaa tacacgttaa ccagtgcact taaggacagc 780
cgatttcccc ccctgacccg agaggagctg cctaaacttt tctgctctgt ctccctcctt 840
actaactttg aggatgccag tgattacctg gactgggagg taggggtcca tgggattcga 900
attgaattca ttaatgaaaa aggtgtcaaa cgcacagcca catatttacc tgaggttgct 960
aaggaacaag actgggatca gatccagaca atagactcct tgctcaggaa aggtggcttt 1020
aaagctccaa ttaccagtga attcagaaaa acgatcaaac tcaccaggta ccgaagtgag 1080
aaggtgacaa tcagttacgc agagtatatt gcttcccgac agcactgttt ccagaacggc 1140
actcttcatg ccccgcccct ctacaatcat tactcctgac acacggctgc atgaccagtc 1200
ccaccccccc gtggccacca atggctatga catcattgga gcagatgcct cctcttcctg 1260
gtccagttac ttctattact gcaccatttt atgatgctag cttccgttgc caagtctgcc 1320
tcccgctgac tgagggaggg tagggttacc ttgaacgaaa cagaacttga ggggcccaag 1380
ccttatctca gcctttcctc aatatggggt tccgttggat tggggctcct ccatgactag 1440
tgggaattac tgtgggttca gaagaccctt gtctggtatt tgccacatgg ggtattggcc 1500

acacgctgga agctgaaatt gatgatccct gaaggttgaa cccacacaca cccctgcagc 1560
ctccccagat gaagtaggtg tattcccctg gcagtctggg caacggagac caagaaacat 1620
ttttaggttg ttttaaattc cttttttaa acttccagtt tattgcgtac caagagttga 1680
ttacaacctc catgcttcat aagcggacgc cacgttaggg ttggacgtgg gcaccacgag 1740
tcctttgagg ctcctggaca gagacccaca tcaagatcgg aagccctttg ggtggcgttg 1800
cagatctcat tgctcagtag gccgtgaaga ttttcatcct catcccactc tcagttggat 1860
tttctggcac tcttcctgca ttgagtctcc tgattactga acagagctcc gtcatgtagc 1920
ctgctgagga atggaatgga atggagatgc ccacaggagg tcctgatgtc atcactgcac 1980
gcaggtgtga gaggagagac ctcttctgca ccgcctggct acctcactcc tctgctggta 2040
gcagtgccta tagctggacc taagttctca gaagcgtaga tgtgcaaaca agcgattgag 2100
ttgggcttta ggaggacaca tcataggaga gaatccaggg tctgtaagct ggttttcttt 2160
tcaggtgaca tcctgagggg cctgtaagca ggggagctcc tttttctagt ttgcctgtag 2220
aggtgggaag actgttggtg tttctgtcct ttacaggaca ttaggaaaca gttgtgtaat 2280
tacacaaggt ggacctttat cttgcctgac atgctgggaa tcttcacccc accagggcaa 2340
atttccaaat agctcatttt attctaggtc tttcaaactt tcatgtgaca tatttccctt 2400
tcccattgtt gctgatttcc aaatcgctgt cagcaatttt ttcctctctc cttgcctatt 2460
cttcactcat ttggtggcaa agttcataga actaggggac ttggaagatg ctttgaaaat 2520
attgttacaa aggcactgct aaaatgattc acagggagag tggccagttg gaagaaggat 2580
cctaaggatg tgacactggt tttcaacaac atgcttagag aactcatgaa gtggattggg 2640
tgtcaaccca gtgaacatgt ttttatttaa tttatttttt gaagtttatg tggtgatggt 2700
gtggctttcc gaaatgggca aatattcaga agatcttttg cattttcttc tgccaggaat 2760
ggggaagggg agtgggggca caatctgaga aaggacacct gtgctgttct aggcatcgct 2820
ggcaagtttg tgggaaggga tgggcaaggg tgagtgggtt tgctccacac cgtcctgtgc 2880
tgctcgagag gacctgggac gtgcgaggga aacgtgggtg acggtgccta ggctgcggcc 2940
cttcactgct gtgctgggtt cctgcagcct gctacgtttc ccttggcaat gtaaatgaag 3000
atggaggggt cgtttcgtga tttcctgctg ctgagaataa atgtcttgtt aaaaacgtgg 3060
caacggttac tcttaggtgt catggatcga tgtcagggtg gtcagctctg gactaagcca 3120
cccacctcca atttgtacaa cagtattgat acatagggct acactcatta ctgttcaagt 3180
gttctatgtt aagagttgtg tttaatttct aaagattaaa aaaagcaaaa aaattggtgc 3240
taaaccttca cccctgagca cgctcagtga gactggtcat gcaagcattt acagtgccat 3300
gctcctcaag accattggca gg 3322

<210> 100
<211> 3621
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8227004CB1

<400> 100
cggtgagcgc gtttcccgcc tgagcgcaac tagcggcggg tcgtgggcac ctccagacca 60
gcatgttttg gaaatttgat cttcactcat catcccacat agacacactt ctagaaagag 120
aagatgtaac actgaaggag ttaatggatg aggaagatgt tttacaggaa tgtaaagctc 180
agaaccgcaa acttatagag tttctgttaa aagcagaatg tctcgaagat ttagtctcat 240
tcattataga agaaccacct caagacatgg atgaaaagat cagatacaag tatccaaata 300
tatcttgtga gttgctcact tctgatgtct cccagatgaa tgatagactg ggagaagatg 360
aatccttgct aatgaaatta tatagcttcc tcctaaacga ttcccctttg aatccactac 420
ttgccagttt cttcagcaag gtgctaagta ttcttatcag cagaaaacca gaacagattg 480
tggatttctt aaagaagaag catgattttg tagaccttat tataaagcac ataggaactt 540
ctgctatcat ggatttgttg ctcaggctcc tgacgtgtat cgaacctcca cagcccaggc 600
aagatgtgct gaattggtta aatgaggaga aaattatcca gaggcttgtg gaaatagttc 660
atccatcgca agaagaagat cgacattcaa atgcatcaca atcactttgt gaaattgttc 720
gcctgagcag agaccagatg ttacaaattc agaacagtac agagcccgac cccctgcttg 780
ccactctaga aaagcaagaa attatagagc agcttctatc aaatattttc cacaaggaga 840
aaaatgagtc agccatagtc agtgcaatcc agatattgct gactttactt gagacacgac 900
gaccaacatt tgaaggccat atagagatct gcccaccagg catgagccat tcagcttgtt 960
cagtaaacaa gagtgttcta gaagccatca gaggaagact tggatctttt catgaactcc 1020
tgctggagcc acccaagaaa agtgtgatga agaccacatg gggtgtgctg gatcctcctg 1080
tggggaatac ccggttgaat gtcattaggt tgatatccag cctgcttcaa accaatacca 1140
gcagtataaa tggggacctt atggagctga atagcattgg agtcatattg aacatgttct 1200
tcaagtatac atggaataac tttttgcata cacaagtgga aatttgtatt gcactgattc 1260

ttgcaagtcc ttttgaaaac acagaaaatg ccacaattac cgatcaagac tccactggtg 1320
ataatttgtt attaaaacat cttttccaaa aatgtcaatt aatagaacga atacttgaag 1380
cctgggaaat gaatgagaag aaacaggctg agggaggaag acggcatggt tacatgggac 1440
acctaacgag gatagctaac tgtatcgtgc acagcactga caagggcccc aacagtgcat 1500
tagtgcagca gcttatcaaa gatcttcccg acgaagtcag ggaacgatgg gagacgttct 1560
gcacaagctc cttaggagaa actaacaaga ggaacacggt agatctaatg caacaaatga 1620
cgtccaattt tattgaccag tttggcttca acgatgagaa gtttgcagat caagatgaca 1680
ttggcaatgt ttcttttgat cgagtatcag acatcaactt tactctcaat acaaatgaaa 1740
gtggaaatat tgccttgttt gaagcatgtt gtaaggaaag aatacaacag tttgatgatg 1800
gtggctctga tgaggaagat atatgggagg aaaagcacat cgcattcaca ccagaatccc 1860
aaagacgatc cagctcgggg agtacagaca gtgaggaaag tacagactct gaagaagaag 1920
atggagcaaa gcaagacttg tttgaaccca gcagtgccaa cacggaggat aaaatggagg 1980
tggacctgag tgaaccaccc aactggtcag ctaactttga tgtcccaatg gaaacaaccc 2040
acggtgctcc attggattct gtgggatctg atgtctggag cacagaggag ccgatgccaa 2100
ctaaagagac gggctgggct tctttttcag agttcacgtc ttccctgagc acaaaagatt 2160
ctttaaggag taattctcca gtggaaatgg aaaccagcac tgaacccatg gaccctctga 2220
ctcccagtgc ggctgccctg gcagtgcagc cagaagcggc aggcagtgtg gccatggaag 2280
ccagctctga cggagaggag gatgcagaaa gtacagacaa ggtaactgag acagtgatga 2340
atggcggcat gaaggaaacg ctcagcctca ctgtagatgc caagacagag actgcggtct 2400
tcaaaagtga ggaagggaaa ctgtctacct ctcaagatgc tgcttgtaaa gacgcagagg 2460
agtgtcccga gactgcagag gcgaagtgcg cggcgcccag gcctcccagc agcagtcccg 2520
agcagaggac tggccaacca agcgcaccag gtgacacttc agtgaatggc cctgtatgac 2580
gggtgacgtc tgctgctgct gactgaggac tgcagaccgc caccactcag ggctctgga 2640
ggggtcagct ggagcccacc aagctgtcac tgctgcactc actctgcaag ggatcaggac 2700
cagcaacctt tatattctag attctaagac attgtacaga gaaattcaga agtgtaaaaa 2760
tattgcacat tgacaaatac caagaatttt tgcgtatgtt tatattgtat tgttctaaat 2820
aatgggtagc ctgtgaaata agatcttgcc acccatgtaa taatagtagt aatactatag 2880
ttaaaatggc tgtaagaata gttttataaa agtgaataca cagatctatt gtatttgaaa 2940
cataactttg acaattatta gtgtgaccaa agtattaggc ggttttcata catttttcac 3000
cttgtacaaa attatgaatt catttttcct ccaggccgac aaggagttgt agaatgaaaa 3060
tgccctctaa gtgttatttt ggttgttcta acttacaaaa gtgattttga ataagaaata 3120
tttggtgttc tttttataac cagtttttga ttggtaattg ttttctgtat tgtttaaaac 3180
ggatcaaaaa tgtaagtcta ttggtagaga ttaagtattt attgctacat catagttgat 3240
aaattgatgt tatcgtaaag ccatatgttc tgttcaagtc ttgtttgctt gaaatgatta 3300
ttcctacaag tgaaacacta gactatttgg agtgtatatg gcttgtgttt tgggattttt 3360
tttttttttt ttggcttttg tttttgtttg tatttttggt tcgtttggta gttcatctgc 3420
cttttaaccc attcaccaaa atttaccttg ttaacaagca tcaccaatga acatttcaga 3480
gcaatctgca tatttaacag acctaaaata aatcctatta agcaagtcag ttgaaaatgc 3540
tcgtgctgct aatgggaata aagtgcgttc catttacagg gcaagtattt ttaaaatagg 3600
aatcaaaatc tgggcaccga a 3621

<210> 101
<211> 2704
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3044763CB1

<400> 101
ctggctcagg gggcggctgc ctcgccaggc gccacgaaga cccggcgtcc cacaggtccc 60
ggaggccgct gcgtcgcggg gagctgggct cgggccgccg gcggttgcgc cggggcagcc 120
ccgcagggag agcaccatga ggactgactc gggggctcgg ctggaggagg gccacctgcg 180
tcctccacga gctcttccac ctgtgccaag tcaagatgat attccactta gtcgtcctaa 240
gaaaaagaag cccagaacaa aaaacacacc agcaagtgct tctttggaag gccttgctca 300
gactgctggt cgaaggccct ctgagggcaa tgagccatca actaaagaac tcaaagagca 360
cccagaggct cctgttcaaa gaagacagaa aaagacaagg ctacctcttg aattggagac 420
ttcctccacc caaaagaagt catctagttc atctttatta cgaaatgaaa atggtattga 480
tgcggagcca gctgaggagg cagttattca aaaacctcgg aggaagacaa agaaaaccca 540
gccagcagaa ttacagtatg ccaatgagct aggagtagaa gatgaagaca taatcactga 600
tgagcaaact actgtggaac agcagtctgt attcactgca cccactggca ttagccagcc 660
tgtaggcaaa gtatttgtgg aaaaaagccg gagattccag gctgctgatc gttcagagtt 720

gataaagacc acagaaaaca tagatgtgtc aatggacgtg aagccttcct ggaccaccag 780
agatgtggca cttacagtgc accgggcttt caggatgatt ggtctctttt ctcatggatt 840
cttggctggc tgtgctgtgt ggaatattgt tgtgatatat gttctagcag gagatcagct 900
atccaacctc tcaaaccttc tgcaacaata caagacccta gcgtatccat tccagagtct 960
tctgtacttg cttttggctc tgagtacaat ttcagctttt gacaggattg actttgctaa 1020
aatatcagta gcaatccgaa attttttggc cctggatcct acagctttag catctttctt 1080
gtactttact gctcttatac tatctctgag tcagcaaatg acaagtgaca gaatccacct 1140
ttacacacct tcttctgtta atggtagcct ctgggaagca ggaattgagg aacagattct 1200
ccagccatgg attgtggtga atctcgtggt ggctcttctg gttggattat cttggctatt 1260
tttgtcttat aggccaggca tggatcttag tgaagagtta atgttctcct cagaggtgga 1320
agaatatcct gataaagaga aagaaatcaa agcctcttca taataccagc tcacctttgg 1380
aggaataatg acccagtatt tgtcccattt ttgtttttaa atgtattttt ataagatgta 1440
tacatgtgta tttgtaatag atttttttgat tatataatac tgaaacatct ctcaatatta 1500
tgaaaaatgt taaaattgtg tttgctgttt atgtctaaac attaatttgt ctagcattat 1560
catcttaatg acaaagggaa taatgaacta gaaaccagca agtgaaagtg tttatttcct 1620
attttctcaa aacagttgta tttataacta ttaccttaaa aagcactggt ttagaaaaag 1680
ccataactta aatagtgtta taaaatatat atcaggttta aacataaatt tagcgaatat 1740
ggtagaaggg aaaaaagcct tcatttttga cctcccctta ctgaataaat tgaaatatga 1800
agtttgtctt ttctgaaact ggcttagtga ttgagtatca tgtaataatt ataatataat 1860
ttagcttgaa agatgctcta cattatgacc aataaaaagg gaatgtatgt tttgtttgaa 1920
aaattatata acctaaattt cttacctaga agtaactaaa aagttgcttc tcattataat 1980
ctgtactagt ggttctcata tctggttgta catcaggatc atagagggag atttaaaaa 2040
tctatgtagc aaaagaggct gagcagggca aggctcaggg gaataaagga gagacctgtg 2100
agcttgtggg ctcccagttg acatctgcag taccctttcc tctgtccctt ttttctgagt 2160
gacaagaaat aggagtagaa attcaccatc tctgttctcc agctctttgt cagagaggtt 2220
tctggttttc caggaaatcc ccctcttgag atgggtcctg cactgggaaa gatctcttca 2280
gaactatatc caatggtacc tcagctgttg tatatgagac cctaaagctc aagttgagga 2340
gaagactagc tatgggaaaa aatgtttcaa aggctgttga gcacatcaat aaaactattg 2400
agcctgccct gattagcaag cacctgaatg ttatagaaca gaagaggatt gacaaattga 2460
tgatagagac agttgaccct gacaataggt ctaaatttgg agtgaacatt atactgggaa 2520
tctcttttgc tgtttgtaag gctggagctg ccgaaaaggg attctccctg ctgtcacaga 2580
attgtgaatt tgctggcaat tctgaaggca tcctgctagt tccagctttc actgtgacca 2640
gcaatggttc tcaatctggc aataagctgg cagtatagga gttcataatc ttccccgtca 2700
gcaa 2704

<210> 102
<211> 1345
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4044519CB1

<400> 102
cagaattcaa attcaatctt gtcacatccc tgtttaaaac tttttggtta ctacttactg 60
cttttaggag aaagaacaaa atcctcaggt ggctgacaga tctgccatga tctggttccc 120
acctacccct tggctgggtt ccacaccatt tcttgtctct ctcagtgaca ggacagttct 180
gctaattttc tttgagttaa ctaaatgcct aattccaccc actcccagta gcccttcatt 240
tcaaaggatt cacatgggct gctccttctg cttgaatgtg cttccttttt tttcttttgt 300
tcttcactat ggtagcttct acctgccctt cagatctcag gctcaaagat tccttcctca 360
agaacatggt ccctgcattg aaaggctgtt tcaggactta ttttatatgc ttccttctca 420
ttctaatttt ccaattgaac ccaagttcca gtttaccatc cagtttacca gtatatcttt 480
tttcttttct ttcttttttt tttttctttt tttttcttga ggcagagtcc tgccctgtca 540
cccaggccga agtgcagtgg tatgatcata gctcactgta gtctcaaact cctgggctta 600
agtgatcctt cagcctcagc ctcccaagaa tctggaacta caggtgtgca ccaccatacc 660
tggctagctt atctaatttt tttgtatagc cagagcctca gtatgttgcc catgttggtc 720
tcaaacttct ggactcaagt gatagtccct ccttggcctc tcaaagtgct aggattacag 780
acctgagcta ctgcacctga ctactgtttt cccttttttaa atctatttat taatgcctat 840
ctctcctatg agactatcag cgcacacata catctttttct atattcttca taactttgtt 900
cttataactt cctaccacat gacctgtcac acaatagcca ttcaaaaaac ttgttcaatg 960
aatgcacaca tagattattt gcttctccga atatgccaca aatgcatgtt gatataatct 1020
tctgcaaatt gagttccgtg ttcgtatata aagtgtatgt actttctaca tatgttgaat 1080

atatttagat tatcttcctc attcacaaat gaagaaacta agatctaagg gaggtaagtg 1140
acttttccag ggtcactcag ggaatttgtg aacagatcgt aaccaaaatt caggccagta 1200
aactgacagt ccagacagtt tagtaaaatt gaaactgcaa atggcgattg tcttcaactg 1260
acttgaatcc aaatcctgaa gttatttctt gttaagtcta ttactttggg aaatgttccc 1320
tactgaagaa aaaaaaaaaa aaagg 1345

<210> 103
<211> 1607
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 71351918CB1

<400> 103
ttgggcggaa gaggtgggct ggtggaggcg gggtcgagat ggcggcgcct ttgaggattc 60
agagcgactg ggcgcaagcc ctcaggaagg atgaaggggа ggcctggctg agctgtcatc 120
ccccagggaa accatctttg tatggcagcc tgacttgtca aggaattggc ctagatggca 180
tcccagaggt tacagcttca gaaggattta ctgtgaatga aataaacaag aaaagcattc 240
atatttcatg tccaaaggaa aatgcatctt ctaagttttt ggcaccatat actacttttt 300
ccagaattca tacaaagagt ataacatgcc tggacatttc cagcagagga ggtcttggtg 360
tgtcttctag tactgacggg accatgaaaa tctggcaggc ttccaatgga gaactcaggg 420
tttttcccat caggccttgt ggtcctgagt gggggaatgg atgcccagct gaagatatgg 480
tcagctgaag atgctagctg cgtggtgacc ttcaaaggtc acaaaggagg tatcctggat 540
acagccatcg ttgatcgggg gaggaatgtg gtgtctgctt ctcgagatgg gacagcacga 600
ctttgggatt gtgggcgctc aggctgcttg ggagtccttg cagattgtgg ttcttctatc 660
aatggagtgg cggtgggtgc tgctgacaac tccataaacc ttggctcccc tgagcagatg 720
cccagtgaac gggaggttgg aacagaggcc aaaatgctgc tcttggcccg ggaagataag 780
aaacttcagt gcttgggact acagagcagg cagctggtgt tcctctttat tggctcagac 840
gctttcaact gctgtacttt tctctctggc ttcttgctat tggctgggac tcaagatgga 900
aacatttatc agctggatgt gaggagtcca agggctccgg tacaagtcat ccacagatca 960
ggagcaccag ttctatccct gctaagtgtc agagatggat tcattgctag ccaaggtgat 1020
ggaagctgtt ttattgtcca gcaagactta gactatgtca ctgagctcac tggggctgac 1080
tgtgaccctg tgtacaaggt agccacatgg gagaagcaga tctacacatg ctgtcgagac 1140
ggtcttgtac gacgctacca gctttctgac ctctgacttc ttggaaagag cagtcccggt 1200
tagtgaaaag gtttgaccct gatcaacaat gagcagaaac atcatcagtc cttcccaagg 1260
accatggcgt ttaatgtctt gggcacccct tggaaatcac agaaagtcag ctgtactggc 1320
cgtgtggaac tctcatccca agacctactt tgaactgagt aagaaggtca ttgtgcccac 1380
tgcatttgtt ccaacttctc cttgtataaa ctcaccccag caacacaggg caaggatata 1440
gatgctttta gtttgttctt aaaccagttt tgttaaatgt ttacaaggac ctcagtacta 1500
aagcctgttc tctggaggaa ataaagaaaa tatgtttgga ggtgcctgaa tatgaagaac 1560
tttgttaaat actcttactc cagcaaaaag ttggtcagga cgatgga 1607

<210> 104
<211> 2622
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8109363CB1

<400> 104
atcaattccc tggtcccagg agaagtttta ttgtgagttt gttaatttaa gaactcactc 60
agcagaagga catgtaatgt ctgagtaatc ttgatggaag aaaggaaata gatctgtgat 120
taaaacacac acacacacac acacacacac accctacttc ggagagagaa tgtagaatgg 180
aggaggtgag gaatttaact aatagagatg ctttttctaa caggcagaaa ttagagaaat 240
ggatcaaaag atttaaacat tgattacata taaacgagtt tcctttttttt ctacatagca 300
cacaggtgct caacatttaa tgaataatat aacagatact gtcttaaaag tttaaaaaca 360
ggatttgtac ttaatcctaa attcctctta tcacacacca gaaaataatt tggattttat 420
ttacaaacta catccaaata ctcttaagga aagagtgtca tttagcttgg agtaggagact 480
gactggggtc taagtgttgg aaattaaatt ctgccattca gtccaacctc aaaggtcttc 540

caggcaagga gttttgcaag gactagaaac cctttcacac ttgaagtgca ggtgttggct 600
gaaagaatag agggacattt cacaaacaca gaacacggct cccctacagg aaagggatgg 660
agacagctac ctttgctctt gctctcacat tctctgtcct catggacttt cttctctgtt 720
tccctctgcc tatctatctg cctggttctc ttgcttctcc tggccaacca gctttctcat 780
tagtgcatat gtagctgaaa aatccacccc tcaagaaccc actccacaca acctttcctt 840
ttcagtaccc accacgctct gactagcatc tctgtgatgc ctaaggcaaa tttttgagga 900
tctgattggt ttagttcatc tgcctatggg tgagccaccc tggggtcagg tgcctgtcca 960
atcacctgtt accaaggagg cagggccgtg tcatttaagc atggccccat ggactgaccc 1020
tttcagcagg gacttttatt gggggctggt acccagaaga aggagcatga gtgggcaaga 1080
ccttcaaaaa gaacctacta cacactggaa ctgactagta ctcctgggat gccatgccac 1140
gcagcagcca gcactcggag agcagtcccc tggatactac cacacagaga aaaggggcat 1200
cttcacttgc acatcaggtg agggtgcaca cacttgaaac tctgctagac tggccagagc 1260
tgccccagcc acttctcacc ccaccaccag ttattgatac agccgctggg tctaggaagc 1320
gatttctgaa caaggcccag ttagcacaat gtcttgctca gcagacaata aatacttgta 1380
aattgaattg catgatctta gctcaggtgc tactaatgtg gctcacggcc acacacttgc 1440
atggataaga aaaccaatga gcccctcaac agagtctttt gaaacaaaat tggtttcaag 1500
tgaaacttac attatttttt ggctgtgacc ttctgaatat aataacttag tgtcctgaga 1560
attgaaaggc ctcttgattt cttaggtgaa ctgttacaga atcctagcct gtggtaaatg 1620
ctaaaatgta aggcgtttct gagaaggaat atttaggaaa tataattgtg ccacacacac 1680
attactttag aagatatgca gttgctctga ttctgcaggt ttaaaaaaaa tgggaagttt 1740
gtgatatggc agaatctgtc tccattggtc aaaccatact tcagtgagag ccattataag 1800
tgaaatctaa aaattgatcg ttttcattct ctgactttgg tcagactttt gcaaattatg 1860
gagagaggcc atcaacatat tttacataaa ttactgtggc cccatcttat ttctcctcat 1920
ttcttcctta tgtttaaaca cgccccccatc tcacacacaa actgagaaaa gatttgtatc 1980
aaacagaatc agaggcatat gaatggactg gaaattgtga aaaaggtata aatatcctag 2040
accttgatca actcacttta aactggtctc tgattcctca tctgtcaaat gagtataaga 2100
ccctctgtac ttcccatggt gattttgagg agcacatgca taatgtatct gaaaatggct 2160
tgacaagcac ttgaagccat tcttactaaa ttattagttt tgtaattcgg ttttatccag 2220
gaaaatattt tagctcctcc tatctgctga attaatttat acagtcacat gttccatgac 2280
accatgtttt gccaaggaaa aataaagaaa ttcctaaatt tgaaattgcc cacaagatct 2340
agcctcatct ggagttgctg ttattgtcct tgttgtctat gatttggagg atatgtaagt 2400
acttgtacat gtgctttaac tctgggtgaa aaaccccaac cggtgagtct ttaaaaccat 2460
attagacaca ttgggcaaag tactgagcag aagctttgaa tgtagctgtt gttattttaa 2520
tgttcgcttt ctaatgaact ttgtacgtcc ttagcctctg ccgggaacag agtgttcata 2580
tgctgttaga atttttattt gtaaaataaa atgacaaagg ct 2622

<210> 105
<211> 3489
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1272746CB1

<400> 105
cggcagtcgg gaccgactgc aaggcaggtt gagatgatgg atctgaggct tgcagctcgc 60
agtcccaacg agccttaggg accagggcta ccaaaagggt ggggagaacg gtggtggcat 120
ctgtctgttg cctgggatcc aggcgcccgg ttggggaatc cttacttcag gaagcgcagc 180
cgcggctctg tgttttgaga gtgcaaatgt catttgtcag agtgaaccgc tgtggtcccc 240
gagttggtgt aagaaagaca ccgaaagtaa agaagaagaa aacttcagtg aaacaagaat 300
gggataatac cgtgactgat ctaaccgttc atcgggcaac tcccgaagat ctggtacgcc 360
gtcatgaaat acacaaatcg aagaatagag cattagtaca ctgggaactc caagaaaaag 420
ctttgaagag aaaatggagg aagcagaaac cagaaacttt aaatcttgag aaaagaagat 480
tgtctatcat gaaggagatt ctttctgatc aataccagat gcaagatgtg ttggagaaat 540
ctgatcatct aatagctgca gcaaaagagc tgtttcctcg taggcgcaca gggtttccaa 600
atgtaacagt ggctcctgat tcctctcagg gtcccattgt ggtaaatcaa gaccctatca 660
cccaatctat ctttaatgag tctgtcatag aacctcaggc tcttaatgat gtagatggtg 720
aagaagaagg aactgttaat agccagtcag gagaaagtga gaatgagaat gagttggata 780
actctctaaa ctctcagtct aacacgaata cagacaggtt tctccaacaa ctaacagaag 840
agaatttga gttaattagt aagttgtgga ctgacattca gcagaaaata gcaacccagt 900
cacaaataac tcctccagga acgccatcat ctgctctttc atcaggggag caaagagctg 960
ctctgaatgc taccaatgct gtcaagagac tccaaaccag gcttcagcct gaagaatcta 1020

```
ctgagactct agactcaagc tacgttgtgg gacacgtgct gaactcaagg aagcaaaaac 1080
agctgttaaa taaagtgaaa aggaaaccga atttgcatgc tctttccaag ccgaagaaaa 1140
acatatcatc aggtagcaca acctctgcag acttaccaaa taggactaat tccaacctgg 1200
atgtcctcaa acacatgata catgaagtgg aacatgaaat ggaagaatat gagcggtgga 1260
caggtcgcga ggtcaagggt ctgcagagca gtcagggtct tacaggcttc actttgtcgc 1320
tggtgagctc cctctgtcgc ctggttcggt accttaaaga gagtgagatc cagctacgta 1380
aagaagtaga gacaaggcaa caactggaac aagtattagg tgatcatcga gagctcattg 1440
atgctctgac agctgaaatt cttcgtctta gagaagaaaa cgctgctaca caggcaagac 1500
ttcagcagta catggtcaca acagatgagc aactgatatc actcacacat gctattaaga 1560
actgtcctgt gataaataac agacaagaaa ttcaggcatc agaaagcgga gccacaggta 1620
gaagagttat ggacagtcca gagcgtccag ttgtaaatgc caatgtctca gtgccattga 1680
tgttcagaga ggaagtggct gaattcccac aggaagagtt gcccgttaaa ctgtctcagg 1740
tgccagaccc tccagataac atgaatctgg ccaagaattt tccagcacat atttttgagc 1800
cagctgtgtt gttaacacca cccaggcaga agagcaactt aaaattctct cctcttcagg 1860
acgtattgag aaggactgtt caaactcgtc ctgctccacg acttcctcca actgtggaaa 1920
taattgagaa ggaacaaaat tgggaagaga agaccttacc tattgataca gacattcaga 1980
attcaagtga agagaatcgt ctcttcactc agagatggag agtctctcac atgggagaag 2040
atttggagaa caaaactcag gctccttttg ttaacctctc acagcccctc tgcaattccc 2100
attccaacac tcaacagtca agaagcccca cattctcaga agagctccca gtactgggag 2160
atgggcagca gctgagaaca aatgagtcat taatacaaag aaaggacata atgacacgaa 2220
ttgctgattt gacattgcag aattcagcta tcaaggcaca tatgaataat attattgagc 2280
ccagaggaga gcaaggggat ggactccggg agttgaacaa acaagaaagt gcaagtgaca 2340
tgacttctac ttttccagta gcacagtctc taacaccagg tagtatggag gaacggattg 2400
cagaattgaa tcgacaaagt atggaggctc gtggaaaact actgcagttg atagagcagc 2460
agaaacttgt tggtttgaat ctttctccac caatgtcacc tgttcagtta cctctcagag 2520
catggactga aggagcaaag aggacaattg aggtatctat tccaggagca gaagccccag 2580
aaagctcaaa atgtagtact gtctctcccg tcagcgggat aaatacaaga agatcttccg 2640
gggctactgg taattcttgt tctccactaa atgccacctc aggaagtggg agattcacac 2700
ctcttaatcc aagagcaaag attgagaaac agaatgaaga aggctggttt gctctttcta 2760
cccatgtatc ataagtgaag tcaagtctca ctgagtttgt tcttaataat ttattacttc 2820
cctttccctg ctctgacttt taagttttta tatccttctt tcagataaaa cctacaaaga 2880
tcctgtgaat tagtactaat gaaacagaga aataaagcaa ttattttttg actttctcaa 2940
ataagttttc aacaaccaac tgacctacag ctccctgtga atgaaacttt gactgtagag 3000
aagttaaagt tttgtatttt aatactttct taagtatttt aagggttttt ttttattaat 3060
acatttttac ctttatcttt attcagggtt ttttgaggtt tagtatatct tagttgatcc 3120
atttatttat ttattttcct gaagaactaa tttgctttgc atgtaactta caataaaatt 3180
cctctgtgtg attagagtct ggctttcagg aatatgtaac acccactttt ctttcttttt 3240
tcttgggaag tcattgctgt tcatcacttt ttccatcaag tgaaatatac agccttaaaa 3300
acaataggcc tgggagtgtt tgtttatatt tgcttaagca ggtaagagtg gttgcattta 3360
ttataaaatc ttatgtagtt tttaaatgtg aaaatgtcat caataatggt aatgcaaatc 3420
aaataaatat gattctagag ctcaaaaaaa aaaaaagggg ggggcgccga ataagtgggt 3480
tcgtgaacg                                                         3489
```

<210> 106
<211> 2269
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1839974CB1

<400> 106

```
ggctggagga ggagagcggc ggcggcggga gcagcgaagg gggcggcagg atcctccagg 60
ctgccggctg ggaaggcgtg ggcgacccgg tgtgtggcgc gcccagagcc ccgcgtttca 120
gccctaggga aggaagccag ttgagggaag ttctccatga atgtacgtca caatgatgat 180
gaccgaccaa atccctctgg aactgccacc attgctgaac ggagaggtag ccatgatgcc 240
ccacttggtg aatggagatg cagctcagca ggttattctc gttcaagtta atccaggtga 300
gactttcaca ataagagcag aggatggaac acttcagtgc attcaaggac ctgctgaagt 360
tcccatgatg tcacccaatg gatccattcc tcccattcat gtgcctccag gttatatctc 420
acaggtgatt gaagatagta ctggagtccg ccgggtggtg gtcacacccc agtctcctga 480
gtgttatccc ccaagctacc cctcagccat gtctccaacc catcatctcc ctccctatct 540
gactcaccat ccacatttta ttcataactc acacacggct tactacccac ctgttaccgg 600
```

acctggagat atgccgcctc agttttttcc ccagcatcat cttccccaca caatatatgg 660
tgagcaagaa attataccat tttatggaat gtcaacctac atcacccgag aagaccagta 720
cagcaagcct ccgcacaaaa aactgaaaga ccgccagatc gatcgccaga accgcctcaa 780
cagccctcct tcttctatct acaaaagcag ctgcacaaca gtatacaatg gctatgggaa 840
gggccatagt ggtggaagtg gcggaggcgg cagcggtagt ggtcccggaa ttaagaaaac 900
agagcgacga gcaagaagca gcccaaagtc gaatgattca gacttgcaag aatatgagtt 960
ggaagtaaag agggtgcaag acattctttc gggaatagag aaaccacagg tttctaatat 1020
tcaggcaaga gcagttgtgt tgtcctgggc tccccctgtt ggactttcct gtggacccca 1080
cagtggtctt tccttcccct acagttacga ggtggcctta tcagacaaag gacgagatgg 1140
aaaatacaag ataatttaca gtggagaaga attagaatgt aacctgaaag atcttagacc 1200
agcaacagat tatcatgtga gggtgtatgc catgtacaat tccgtaaagg gatcctgctc 1260
cgagcctgtt agcttcacca cccacagctg tgcacccgag tgtcctttcc cccctaagct 1320
ggcacatagg agcaaaagtt cactaaccct gcagtggaag gcaccaattg acaacggttc 1380
aaaaatcacc aactaccttt tagagtggga tgaggtaagc ttattttcat attcacccat 1440
ctaaaacacc atttcaaaaa caaaattaac tgaacaaatg ccatcttatt cttttataag 1500
acatgctaaa cacagaaaaa aggaatgcta gtagagaaat ttaaatcaat agttattgac 1560
tatggaaaat gaagttagat tgaagtagta ttatgagttg tgtattgaaa tatatgtacc 1620
tggcttattt gcaagcacta ctgctgaaag aaaaaaagtg gggtcttatc tattttggat 1680
ctccagtgtc tagccagcca ataatttgta ttaaaaattg actgctttgt catcttgact 1740
gctcagatga ttttgataag gtaaaacaca ctgccgcttg gaaagcataa tttgcttttc 1800
ctttttgaat gccatgtatg tactttggtg ttacactgtt gagtggcctc tgtgagagca 1860
gagcagtaca tagaagtcac attctaccca tgtgcgagtg ccctgtgggt tagaatgtcc 1920
tgtatcggat gcccccttg ctgttcatgt caccagaggc ccttggaagg tatctacggt 1980
ggagtgtcct tattctcaaa gccactgtgt gccaactggt ggacccagtg catttctggc 2040
ctagctaact ctacacacag aatctcaggc aggaaacact tgacctagaa gcctttcatt 2100
tctaaaaagc aactgtaggt cgggcgcgat gactcacgcc tgtaatccca gcactttggg 2160
aggccaaggc aggtggaaca cgaggtcagg agatcaagac catcctggct aacacgatga 2220
aacaccgtct ctactaaaaa tacaaaaaaa aaaaaaaaaa aagatcttt 2269

<210> 107
<211> 3075
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1877336CB1

<400> 107
aagaagacca gagccagccg ggtggcacag cggtgtcgtg gccgtgttgc tgatcgcctg 60
ggtggttgtt ggcgtgtccc tgcagcgaag gatcctggtt gacagtgaaa aagcagtctg 120
gctcccgagg tccacccctt ataccccaag gtccagatgg cggccaacgt gggtgatcaa 180
cgtagcacag attggtcttc tcagtacagc atggtggctg gggcaggccg agagaatggc 240
atggagacgc cgatgcacga gaacccggag tgggagaagg cccgtcaggc cctggccagc 300
atcagcaagt caggagctgc cggcggctct gccaagtcca gcagcaatgg gcctgtggcc 360
agtgcacagt acgtgtccca ggcagaagcc tcagctttgc agcagcagca gtactaccag 420
tggtaccagc agtacaacta tgcctacccc tacagctact actatcccat gagcatgtac 480
cagagctatg gctccccttc ccagtatggg atggccggct cctatggctc agccacaccc 540
cagcagccat ccgcacccca acaccaaggg actctgaacc agcccccagt ccccggcatg 600
gatgagagca tgtcctacca ggctccccct cagcagctgc cgtcggctca gccccctcag 660
ccctcaaatc ccccacatgg ggctcacacg ctgaacagtg gccctcagcc tgggacagct 720
ccagccacac agcacagcca ggcggggccc gccacgggcc aggcctatgg gccacacacc 780
tacaccgaac ctgccaagcc caagaagggc caacagctgt ggaaccgcat gaaacccgcc 840
cctgggactg gaggtctcaa gttcaacatc cagaagcgac cctttgctgt taccacccag 900
agctttggct ccaacgcaga gggccagcac agtggttttg gcccccagcc caaccctgag 960
aaagttcaga accacagcgg gtcctctgcc cgggggaacc tgtctgggaa gccggatgac 1020
tggccccagg acatgaaaga gtatgtggag cgctgcttca ccgcctgtga gtcggaggag 1080
gacaaggacc gcacggaaaa gctgctcaag gaggtgctgc aggcgcggct gcaggacggc 1140
tcggcctata ccattgactg gagccgggag cccttgccgg ggctgacccg ggagcctgtg 1200
gctgagagcc ctaagaagaa gcggtgggag gccgctagca gccttcaccc tcctagaggg 1260
gcaggctcgg cgacaagggg cgggggtgcc ccgtcccagc gagggacgcc cggggctggg 1320
ggtgccggtc gagcccgggg caacagcttc accaagtttg gcaaccgcaa cgtcttcatg 1380
aaggacaaca gctcttcttc cagcacagac tcccgctccc gctcctcctc caggtccccg 1440

acgcgccact tccgcagaag tgactcccac tcagactccg acagctccta ctcagggaat 1500
gagtgtcacc ctgtgggccg caggaacccg ccccctaagg gccggggcgg tcgaggggcc 1560
catatggatc ggggccgagg cagggcgcag cgtgggaaga ggcacgatct ggcgcccacc 1620
aagcgcagtc gaaagaagat ggcggcgctg gagtgtgagg acccggagcg agagctgaag 1680
aagcagaagc gggcagcccg cttccagcac ggacactccc gccgcctgcg cctcgagccc 1740
ctggtgctgc agatgagcag cctggagagc agtggggctg accctgactg gcaggagctg 1800
cagatcgtgg gcacctgccc tgacatcacc aagcactacc tgcgcctcac ctgtgccccc 1860
gacccgtcca ccgtgcgccc tgtggcagtt ttgaaaaagt cgctgtgcat ggtcaagtgc 1920
cactggaaag agaagcagga ctacgcgttt gcctgcgagc agatgaagtc gatccggcag 1980
gatctgacgg tgcagggcat ccgcaccgag ttcacggtgg aggtgtacga gacccatgcc 2040
cggatcgcct tggagaaggg tgaccatgaa gagtttaacc agtgccagac gcagctcaag 2100
tcgctgtacg ccgagaactt gcctggcaat gtgggcgagt ttactgccta ccgaatcctc 2160
tactacatct tcaccaagaa ctcgggagac atcaccacgg agctggcata cctcacacga 2220
gaactgaagg cagatccttg cgtggcccac gccttggcat taaggacagc ctgggccctg 2280
ggcaactacc accgcttttt ccggctctac tgccatgcac cctgcatgtc tggctacctc 2340
gtggacaagt ttgcagatcg ggagcgcaag gtcgccctca aggccatgat caaaaccttc 2400
cgccctgcgc tgccagtctc ctacctgcag gccgagctgg ccttcgaggg cgaggccgcc 2460
tgccgggcct tcctagagcc cctgggcctg gcctacacgg gcccggacaa ctccagcatc 2520
gactgccgcc tcagcctggc gcagctgtca gccttctgag cacccagcga ggaggggcgg 2580
gggcaggggc tgcagccccc agcgctgcct ttgcggattc tgtttttgag ccgtggactt 2640
gggttgtaaa tttatttgtg gggagtgcgc tccaggaaga gccaccatcc ctgcccccgt 2700
tttcccttgg ttcagcacag gtaaaacggg ttcccctccc tccctgcctt catggatcac 2760
cagctcacgt catgttgcct tctcttttct ttgtgtgtgt gtttatttaa gttatttttc 2820
ttcctcctct cccttttctt tttggccctc cctccctccc tcttctgcca tgtaactgga 2880
ggatgtgcta tgagtttgca aacagctgga ctgtcaggct gctttttttc cagatgttcc 2940
tcctctgcct ccccttcccc tcctctcccc tccttttcct tccttccttc ctttccttgg 3000
agcactgagc accatttgga agcttgagag aaaccaaaat taaagagaga aagagagagc 3060
gtgcaaaaaa aaaaa 3075

<210> 108
<211> 849
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2321054CB1

<400> 108
cgtccctaga gcctcgccgc cctgggaccc gggctgcttc ctgggagagg attacatttt 60
ctccctgctc cacttgactt tcttccggaa ggagcaggtc tcgcaggatc ggggccaaga 120
ctggaagagt tcagccccat ggctgggggt tatggagtga tgggtgacga tggttctatt 180
gattatactg ttcacgaagc ctggaatgaa gccaccaatg tttacttgat agttatcctt 240
gttagcttcg gtctcttcat gtatgccaaa aggaacaaaa ggagaattat gaggatattc 300
agtgtgccac ctacagagga aactttgtca gagcccaact tttatgacac gataagcaag 360
attcgtttaa gacaacaact ggaaatgtat tccatttcaa gaaagtacga ctatcagcag 420
ccacaaaacc aagctgacag tgtgcaactc tcattggaat gaaacctcag aaaaagagca 480
acagaagtaa ttgtttcaag ctcctgattc tttctactaa atcatgaaca gctttaaaaa 540
catttctgtc tgcataaaat tattttactt gtaacttttc cccaattgtt ctgtgcattg 600
ttttgccttt ttaaattaca tctccaagtg gctcaaaagg ccttgacaca gggaacctgc 660
acatatccag gatatgtgta accagcgatg gtgacttgac cttgccaaga cctgtgattc 720
cttcaggata caatcagtga gaaataaaaa cacatcttgg gaagtgggaa tcctggagtt 780
tatgccattt gcaatattaa aaataaaaa tgcaagttat tatttcaata ataaaaaaaa 840
aaaaaaaaa 849

<210> 109
<211> 2659
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2796034CB1

<400> 109
atatactttt tcatttcttt gtatctacaa tggccaactg gaatttctgc ttgaaacatg 60
gagcccacca taatctcctt tttccagtct tctgatggaa tataatcttc atcttcttca 120
gattcttctt ctacttcact atttgtatca aaatatttac atcgacgtgg cggattattt 180
cctgggcatc ttgagaagca acagattgtg atggatcatc attggaagac tgagtttcat 240
cctcctgacc tgatgaatcc tttatattct cctctttatt ttccccacta cagccactgt 300
tgtcatcatt atcagcatct tcatcatctt caccttcttc ttcctcttct tcctcttcct 360
catcttcttc aggtagtcga acagtactac cataaccata aaggctgaga agttcatgaa 420
ttggcatgtc gccttccctt gcaagatctt ctatttcaga gctgaagttt gtttctcctt 480
ccatcatttc ttcctcttct aatgttcgtt catcatcaaa atcatgaacc agcatgtcag 540
ctgatggatc aaattcatgg tcatctgatg ttgctgaacc tcctggactt gaagattcaa 600
cagatggcag catcacatcc gtcttgtctg cattgaacct caaccagttg gttcttaacc 660
acgtccgaaa ctcagccagg cactgggaga atcgagcgac cacaccatag ccgctgccgc 720
tgctgccgcc gccaccttcg ctgctgccgc caccgcctga agaagcccca tccatccgcc 780
tgcagccggg agactcagag cctgggaggg gcctgaggcc cgggccgccc gcctccttct 840
tccagtcgtc gtccggagct tcctccagct ccaccaaccc tctgccctcc cctccgcctc 900
cgccctcctc ctcggcctcg tgccgccagc caaccgcgcg gccgcttgtt gagccaggct 960
ccaccgtcgc catgacgccc ggcaaacatt ccggagcctc ggcccgagcc gctaacgcag 1020
gagcttgggg gtacagggac ttcagaggcg gccaaaaaaa ggggtggtgc accactcccc 1080
agctggtcgc caccatgcca gtctctccgg caggcagtca caagcagcag aacttcgggt 1140
tgaacaatgc cacacaacca aagaagtcta ttagcttttt tgctacaatg aaagcaactt 1200
cagtgaaagg atatactggt gcaaatcaaa gcagaatggc tgtgtccaaa accgtgctta 1260
ttccacctga actgaaaact gtagaaaaac caaatcccaa tataaagaca acacaggtat 1320
ttgacataaa tggaactgat gttactcccc gacctcttta ccatccagat ccacttactg 1380
gtacagcaaa accaagtaaa ctcttgacat cacaagaagg atcacttgga tcagaattta 1440
tatcttccta tagcctttat cagaatacaa taaatcctag tacgttaggg cagtttacaa 1500
ggtcagtttt aggaagcagt acagtttcta agtcaagtgt atcagcaagt gaatcaatag 1560
cagaagacct ggaagaacca tcctataaac gggaaagatt gactagtttc acagatttgc 1620
aagttataag ggcagcacct gaaaaaattg taacaaaaga agacctggag aagaatatag 1680
agataatact cacagagaca gaaacactga gatttttga cttgcccaca gtcatggtct 1740
ctgtagaatc tgaagaagct gagaaagtaa ctcagagaaa caagaattat gaagtccttt 1800
gtagaaacag attaggcaat gacctatatg ttgaaaggat gatgcagact ttcaatggag 1860
caccaaagaa taaagatgtt caatgtgata aaatcataat ggaagataaa ggcataatgt 1920
ccactgcctg ggatttgtat gattcttaca atgctatgga acttgtatct ttatcagtaa 1980
aacaatctgt ggttgagtca agtagtaaag caaatgtact tcctaaagat caggaccaaa 2040
gattgccagg gagcactaca gaaaaaaata gtgaaactag ttctctaatg gacatagaaa 2100
atgtaattct ggcaaaaatc catgaagatg aggaagacca ctcagatgca atattaaaat 2160
ctgacaaatt tcatcaggac ttattttta tggaacgggt tctgatggaa aatatatttc 2220
agcccaaact tgcagcttat cgtcagcttc ctgttttaaa agaacctgaa cctgaagagc 2280
ctgaagatgt tttagaaagt gcaaaacatg aagaggtaga ggaagaatct aagaaggagg 2340
aggaagaaga aatacatgca gaagaatcaa caatacccgc caacttggaa cgactttggt 2400
ctttttcctg tgacttaacc aaaggcctca atgtgagcag ccttgcctgg aataaaacaa 2460
atccagatct tttggctgtt ggctatgggc actttggatt taaagagcaa aaagaggact 2520
ggcttgctgc tggtcaataa agaatcccat ggtaacccaa acaagaataa gaagcatttt 2580
tatctggctg ttgacaaaat tctcataaaa tatataatta catgtaataa ttaaataaag 2640
tctagaaaat tttatcaaa 2659

<210> 110
<211> 2205
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4413112CB1

<400> 110
ggaacgccag cgaccccagc agcgctgcgg acggtgctgg ccgtggccgc tgcggccccc 60
gtgtccaggt gggccaggac gcagcctctg ggcgccgtcg cttttccagc atcgcagagg 120
caaaagcgtg gcagtgggac ccaaaaggta ggactgaggc tctagaactt gcacctgtgc 180
agggactgca aaccagacct gggaggaccc tttcagcagc ccccactcca ccctatccca 240
ggacttccca gcgacccgcc gttctgggag ataccgggag cgtgatcagg gggcggggcc 300
gtttccaagg caaccgctta tttgcatagg gtcccgtcct ggccaacgag ggcgccccaa 360

atgttcagga catagaagaa ggggttaact ggcccggatc tcctcctcgc cttccaagcc 420
cgctaagcac tggggttatc tacccattcc ccagaagggg agactgaggc agcccaccag 480
ccaaaggagg cgaccagact ggggctgcgt tttaccattt cagaagcggc ttgagctggt 540
ctgagctata ataataaaca ctggcggtgg aggcgagggc gaccacaggg ctgaggtcag 600
ggctaggatt ccggtgtctc tacgtaggtt gcttgaatgg ggggcaccct ggcatggacg 660
ctgctgttgc cgctgctgct gcgggagtca gacagcctag aaccgtcgtg caccgtgtcc 720
tccgcggatg tggactggaa cgcggagttc agtgccacgt gcctgaattt cagtggcctc 780
agcctgagcc tgcctcacaa ccagtctctg cgggccagca acgtgattct ccttgacctg 840
tctgggaacg gcctgcgaga gcttccagtg accttctttg cccacctgca gaagcttgag 900
gtcctgaacg tgctacgcaa ccccttgtct cgtgtggatg ggcgctggc cgcccgctgt 960
gaccttgacc tgcaggccga ctgcaactgt gccctggagt cctggcacga catccgccga 1020
gacaactgct ctggccagaa gcctctgctc tgctgggaca caaccagctc ccagcacaac 1080
ctctctgcct tcctggaggt cagctgcgcc cctggcctgg cctctgcaac tatcggggca 1140
gtggtggtca gcgggtgcct gcttcttgga cttgccatcg ctggccctgt gctggcctgg 1200
agactctggc gatgccgagt ggccagaagc cgggagctga acaaaccctg ggctgctcag 1260
gatgggccca agcccggttt aggcttgcag ccacggtacg gcagccggag cgcccccaag 1320
ccccaagtgg ccgtgccatc ctgcccctcc actcccgact atgagaacat gtttgtgggc 1380
cagccagcag ccgagcacca gtgggatgaa caaggggctc acccttcaga ggacaatgac 1440
ttttacatca actacaagga catcgacctg gcttcccagc ctgtctactg taacctgcag 1500
tcactgggcc aggcctcaat ggatgaagag gagtacgtga tccccgggca ctgagcctaa 1560
gatgtcctaa cctccaccca gaacccсttc agtccctgct gggtgactca gggcgtccta 1620
actcctccat ggcctcagtt tccccatctg aagaatgggg acaggaaagg attgtccttg 1680
aggccccagg aagctctgcc gccccctccc tgtccctcat gccgctcctc agctccctca 1740
gctcctagag ggggaagagg agagaccccc aacaagggga caggagggtc actgtgccaa 1800
tcctgtcatc accctcctgt ggatgtacag gcagtgctca ataaatgctt cgaggctgat 1860
gaggctgctg gctcagggtg cgtgggttcc tcaaggtggg gatttctgag ttctaagacc 1920
aagtctccat ctgagactcc caaattgctc cccacctccc atccctgttt ttttttgttg 1980
ttgttgtttg tttgtttgtt tttgaaactg agtctcactc tgtcacccag gctggagtgc 2040
aatgctgcga tctcagctca ctgcaacctc cgcctcctgg gttcaagtga ttctcctgcc 2100
tcagcctcct gagtagctgg gattacagca cccaccacca tgccgagcta atttttgtat 2160
ttataataga gatggggttt cgccatgttg gccaggctgg tctcg 2205

<210> 111
<211> 3042
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7654832CB1

<400> 111
accagctgat attttctttc ctcaggacca gctgttcagg agcatcccgg acgagagcca 60
ggcacatttg agacttggat ccaactaaag accgccgcag attcttctgc agcaatgtcg 120
gtgttagaag aaaatcggcc gtttgctcaa caattatcca atgtctactt tacaatactt 180
tcgctgttct gttttaagct ttttgtgaaa atcagccttg ccatcctcag tcatttctac 240
atagtgaaag gcaaccgcaa ggaagcggca aggatagcag ctgaatttta tggagtaacc 300
caaggacaag gttcctgggc agatcgatca ccactacatg aagcagcaag tcaaggtcgc 360
cttcttgctc tgagaacatt attatcacag ggttataatg taaatgcagt aaccttagac 420
catgtcaccc cattgcacga agcctgcctt ggagatcacg tggcatgtgc cagaactctg 480
ctggaagcag gagctaatgt aaatgcaatc acgatagatg gcgtgactcc gttattcaac 540
gcatgctccc aaggcagtcc aagctgtgca gagctgcttc tggagtatgg tgccaaagcc 600
cagctggagt catgtcttcc atccccaacg catgaggccg ccagtaaagg tcaccatgaa 660
tgtcttgaca tcctgatatc ctggggcata gatgttgacc aagaaattcc tcatttggga 720
actcctctct atgtagcttg tatgtcacag caattccatt gcatctggaa gcttctttat 780
gctggtgctg acgtacagaa aggcaaatat tgggatactc cattacatgc tgctgctcaa 840
caatccagca cagaaattgt aaacttactg ctagaatttg gagcagatat caatgccaaa 900
aatacagagc ttctgcgacc tatagatgta gctacgtcta gcagtatggt ggaaaggata 960
ttgcttcaac atgaagctac cccaagctct ctttaccaac tttgccgact ctgtatccga 1020
agctacatag gaaaaccaag attgcacctt atcccacaac tccagctgcc aacgttactg 1080
aagaatttct tacagtatcg ataaaacagt aaagtaattc taaatacctt gaaaatcaaa 1140
atttctattt cttttgctta aggaatagtt catataaaaa tatgctaaag ataggataaa 1200
agtgagtgtg agatcaccca gggaagcagt aaaatatcaa ttttcatttt aagtgtatta 1260

gtactattgt tttatcattt atgctgtctt tggattatag catatttata atatctattt 1320
tttgttattt caatataccc tcttgtttat tcaggaaaaa ataataaagt cactatctta 1380
tgttttcttt tggattggaa agaaataaag ctcctagtaa agttattcat attatattta 1440
aaaaggtcta tgttcttcaa ctaatgggtt aactttttaa ttttggctat tgtaattggc 1500
taacttgtta ctgatattga cataaaatat gaatatggct tagcaaagag caactaatat 1560
gttttctgaa accaggatga tgctcattca ggaaatccac tatgttgaat gtccactggt 1620
cattagaatg aacacttcat aaagataaat tgccccctc cccccactt acccctggca 1680
aaggactcac aagtcacttc tcttccccat agtacatgga cttgttacag aaaagtgata 1740
ggatgaaagg tcaatactgc cctctcaccc tgagatattt agaacattac acaccttcta 1800
aggctatgtt aagactgcct ccttctcagg agtagcccct gatgatctct tatcaaggcg 1860
gttatatgca tattactcca gccttcttcc atttcagaaa ttgtcaatag ctagtgatac 1920
ttggacttaa gaaaatgaac tatatatcta attaacaaac aggccctgga tcatattccg 1980
gggatcatca catcagtcca aggatataaa aaaatggaag attaaaaact caaattgtgg 2040
agtttttaaa gagtgggtgc cacttaacac attggagacc ttaaaaaatg tatatagaga 2100
acttaatcgc actgtttcta attcttttta cacaaaacaa gggcttagag tatactgtat 2160
ttgcaagtaa tagcaagaga gtattgtata aaagaagata ttactaagct agtcctctgc 2220
aagggtaaat actatgttta ttctcttcag ctttaactat acttgagtgc aataaaaaac 2280
ttaataaaat tttatgcctg ttttgtaatt cagggcctat tatgatttta gaaataagat 2340
aatcaacgta caactctttg taaaacaaaa gcaactttgg gtgagtttta tagatttatt 2400
aaattttgat aaagtaaata ttttgaccac tattttcaac aacccagggc ttattttctt 2460
ctgacttctt ggctaacttc cagactatgg gaaaataaat tatgattgta tgtgaccccc 2520
agatcactaa gcttgcattt tggcagaaag aaaaaaaatc agttgtgatt taaaacgttt 2580
gttgagattt tcaaacatag aaataaagct gtagcatatt cactaaacaa tgttttatat 2640
gaggataggg acctttaaat atgtgtttta tagaatatct cacacatttg taatggataa 2700
cgtgcattaa aatggagttc ttattataga ataacctggt taaacagata ttaaagctgc 2760
ttttgggtaa aacaggtttc caccagtatt ttccctgagc tagaggagaa gtcttagatt 2820
ttaagagtaa tacaacccat ctatcttatt agtcttgtgc caataagaga agtaaatatc 2880
aaacattaaa atggtcactt tacaatttca gtatgtctgt accttgtatc ttagcttgag 2940
tatgattgtc ttatcaaagt agtgttctaa ctttaatgta ttgtatgttt gttattatat 3000
taaagtgtga ttattattca ttggaataaa aaaaaaaaaa aa 3042

<210> 112
<211> 2112
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503849CB1

<400> 112
ggcctccggt gtgggatggc cgcggagccg ggcggagctg gcttgcggct cccggggccg 60
gctctccggc cggagacatg gcccgggggc ccggcccgct aggcaggcct cgccccgata 120
cggtcgccat gcccaagaga ggaaagcgac tcaagttccg ggcccacgac gcctgctccg 180
gccgagtgac cgtggcggat tacgccaact cggatccggc ggtcgtgagg tctggacgag 240
tcaagaaagc cgtagccaac gctgttcagc aggaagtaaa atctctttgt ggcttggaag 300
cctctcaggt tcctgcagag gaagctcttt ctggggctgg tgagccctgt gacatcatcg 360
acagcagtga tgagatggat gcccaggagg aaagcatcca tgagagaact gtctccagaa 420
aaaagaaaag caagagacac aaagaagaac tggacggggc tggaggagaa gagtatccca 480
tggatatttg gctattgctg gcctcctata tccgtcctga ggacattgtg aatttttccc 540
tgatttgtaa gaatgcctgg actgtcactt gcactgctgc cttttggacc aggttgtacc 600
gaaggcacta cacgctggat gcttccctgc ctttgcgtct gcgaccagag tcaatggaga 660
agctgcgctg tctccgggct tgtgtgatcc gatctctgta ccatatgtat gagccatttg 720
ctgctcgaat ctccaagaat ccagccattc cagaaagcac ccccagcaca ttaaagaatt 780
ccaaatagca tagcttgtgt gggacactgg agccgttgtg atggcagcag aagtgttttc 840
cccttaaagc caagcccatt aattttgatg gaacagcagg acgtacaggg catgtctgaa 900
gggcaggaca gctggcacgg cggacgaccc acccccttatc ccctgggagt gcttactttt 960
ctggtgcaga aagattgttg ggaacagaca ggaaccaatg tgggaattca acttcaagtt 1020
caaaaaacag tcccctaggt taaagagcaa gtgtacagga ggattgcagc ctccccgttca 1080
gtacgaagat gttcatacca atccagacca ggactgctgc ctactgcagg tcaccaccct 1140
caatttcatc tttattccga ttgtcatggg aatgatattt actctgttta ctatcaatgt 1200
gagcacggac atgcggcatc atcgagtgag actggtgttc caagattccc ctgtccatgg 1260
tggtcggaaa ctgcgcagtg aacagggtgt gcaagtcatc ctggacccag tgcatcagcg 1320

```
ttcggctctt tgactggtgg catcctcagt acccattctc cctgagagcg tagttactgc 1380
ttcccatccc ttgggggcag cctcgagtgt agtccattag taatcagatt ccagtttgga 1440
cagggtggct ggattgtata tctcgttagt aatgtacatg ctcttcaggt tctagggctc 1500
ctgttagggg agggagaaat gttgaatcaa gagggaaaac aactactatg atttataaac 1560
atattttaat gtaaaaattt gcatttaaaa ggagtggccc tgttttctgt gttaaaaccc 1620
catttggtgc tattgagttt gttctttatt cttttatccc agtgaaaatt gttgatcttg 1680
ctgtagggaa aaattaaact ctttgaatct ccaaacaagg aagtttcagc attcccttat 1740
ggatcagagg aaccttagag gcctgaaatt gttgcttcca gtttagctgc ccctcaaatt 1800
caagtgaata ttttcccttc tccctttacc cttctccaga aataaagcag gtgacagggt 1860
tttcagaatc ttaccatatt gacttgtgta tcttttttaa aaaataattt gtgatattta 1920
ttgcatacat tttcttttgg cagttttgga ttttaggtat ttttttgcat gaggctgata 1980
atggtgatga tgactttata cccttggtac tatactagga actttacata gtatctctac 2040
ttctcacagt cttgcaaagt aggtgatatt tattcccatt ttatagtttg agaaaacgga 2100
attaggtgac tg                                                     2112
```